

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2003, 08:39:40 ; Search time 24 Seconds
(without alignments)
218.220 Million cell updates/sec

Title: US-09-770-509-2

Perfect score: 885
Sequence: 1 ASQEGVEFIVANTDQALG.....LLEGVKGVTDLIVRPLINL 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	480	54.2	394	US-09-134-001C-4834	Sequence 4834, Ap
2	472	53.3	335	US-08-987-146-2	Sequence 2, Appl
3	472	53.3	419	US-09-120-426-2	Sequence 2, Appl
4	463	52.3	437	US-08-883-515-4	Sequence 4, Appl
5	429	48.5	433	US-08-883-515-2	Sequence 2, Appl
6	271	30.6	223	US-09-120-426-4	Sequence 4, Appl
7	196	22.1	283	US-08-961-083-124	Sequence 124, Ap
8	90.5	10.2	2285	US-09-308-375-2	Sequence 2, Appl
9	79	8.9	388	US-09-134-001C-4951	Sequence 4951, Ap
10	77	8.7	271	US-09-189-527-11	Sequence 11, Appl
11	76.5	8.6	422	US-09-134-001C-3034	Sequence 3034, Ap
12	74	8.4	327	US-09-134-001C-3477	Sequence 3477, Ap
13	74	8.4	1529	US-09-134-001C-3945	Sequence 3945, Ap
14	73.5	8.3	450	US-08-861-464-2	Sequence 2, Appl
15	73.5	8.3	450	US-08-396-001-2	Sequence 2, Appl
16	73.5	8.3	450	US-09-323-433A-2	Sequence 2, Appl
17	73	8.2	1026	US-08-194-290-7	Sequence 7, Appl
18	73	8.2	1026	US-08-614-377A-7	Sequence 7, Appl
19	73	8.2	1026	US-09-142-648B-7	Sequence 7, Appl
20	71.5	8.1	267	US-08-557-128-4	Sequence 4, Appl
21	71.5	8.1	267	US-09-242-690A-36	Sequence 36, Appl
22	71.5	8.1	1004	US-09-268-347-30	Sequence 30, Appl
23	71	8.0	267	US-09-302-620B-106	Sequence 106, App
24	71	8.0	267	US-08-935-450-4	Sequence 450, App
25	71	8.0	445	US-08-308-872B-4	Sequence 4, Appl
26	71	8.0	1451	US-08-308-872B-4	Sequence 4, Appl
27	70.5	8.0	348	US-09-134-001C-5513	Sequence 5513, Ap

28	70	7.9	396	US-08-861-774E-84	Sequence 84, Appl
29	70	7.9	2647	US-08-583-562B-8	Sequence 8, Appl
30	70	7.9	2647	US-08-779-113-8	Sequence 8, Appl
31	69.5	7.9	355	PCR-US95-03866-32	Sequence 32, Appl
32	69	7.8	345	US-09-134-001C-3774	Sequence 3774, Ap
33	68.5	7.7	270	US-09-013-881-6	Sequence 6, Appl
34	68.5	7.7	459	US-09-491-785-2	Sequence 2, Appl
35	68	7.7	2314	US-09-268-347-49	Sequence 49, Appl
36	67.5	7.6	409	US-08-743-130A-2	Sequence 2, Appl
37	67.5	7.6	3816	US-09-428-517-3	Sequence 3, Appl
38	67	7.6	458	US-08-618-485B-1	Sequence 1, Appl
39	67	7.6	458	5177002-1	Patent No. 5177002
40	67	7.6	458	US-08-222-619-5	Sequence 5, Appl
41	67	7.6	474	US-08-222-619-5	Sequence 5, Appl
42	67	7.6	474	PCR-US95-04075-5	Sequence 5, Appl
43	67	7.6	568	US-07-732-242C-3	Sequence 3, Appl
44	67	7.6	641	US-08-836-567-10	Sequence 10, Appl
45	67	7.6	3739	US-09-320-878-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-134-001C-4834
Sequence 4834, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4834
LENGTH: 394
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4834

Query Match          54.2% Score 480; DB 4; Length 394;
Best Local Similarity 54.3% Pred. No. 2.8e-47;
Matches 95; Conservative 33; Mismatches 47; Indels 0; Gaps 0;

QY 4 LEGVEFIVANTDQALGRLAPHKITIGKDTKGLGSGKPELGRSAEQKVDIQRMLQ 63
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35 MNVEFIVANTDQALGRLAPHKITIGKDTKGLGSGKPELGRSAEQKVDIQRMLQ 94
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 64 DSMMLFTGSGGGTCTGAAPVAVSARELILTVGVSTPFRSEGPRTLANAGVREL 123
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 95 GADNVEFIVANTDQALGRLAPHKITIGKDTKGLGSGKPELGRSAEQKVDIQRMLQ 154
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 124 AKVVDLIVPNCNLALAKSTTMEAFRADYVLEGVGVVDLIVRPLINL 178
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 KAAVDLIVPNCNLALAKSTTMEAFRADYVLEGVGVVDLIVRPLINL 209
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
US-08-987-146-2
Sequence 2, Application US/08987146
Patent No. 6350866
GENERAL INFORMATION:
APPLICANT: Skatrud, Paul L.
APPLICANT: Peery, Robert B.
APPLICANT: Rokey, Pamela K.
APPLICANT: Wang, O. May
APPLICANT: Rostock Jr., Paul R.
TITLE OF INVENTION: Staphylococcus pneumoniae Gene Sequence
```

TITLE OF INVENTION: FLSZ
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/987,146
FILING DATE: December 8, 1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-11755
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 335 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-987-146-2

Query Match 53.3%; Score 472; DB 4; Length 335;
Best Local Similarity 52.6%; Pred. No. 1.8e-46;
Matches 91; Conservative 35; Mismatches 47; Indels 0; Gaps 0;
QY 6 GVEFIYANTDCCALGRSLAPHKITLTKGDKITGKAGSKPELGKRSADQKVIDIOMLQDS 65
DB 38 GVEFIYANTDCCALGRSLAPHKITLTKGDKITGKAGSKPELGKRSADQKVIDIOMLQDS 97
QY 66 NMLFTGGMGGCTGGAAPVAVASVARELGITLVGVSTPFRSGPNRTRLANAGVELAK 125
DB 98 DMVFITAGMGSGGTGAAPVAVARIKADLGLTGVVTRPFGFSGKRGQFAVEGINQJRE 157
QY 126 YVDLIVVNNLALADKSTMLFAFRYADVILEGVKGVTDLIYRPGILNL 178
DB 158 HVDLIIISNNLLEIVDKITLLEALSDADNVLROGVGIDTLITNPGILNL 210

SEQUENT 3
Sequence 2, Application US/09120426
Patent No. 6197300
GENERAL INFORMATION:
APPLICANT: Fueyo, Joanna Lynn
APPLICANT: Lonetto, Michael A.
APPLICANT: Pearce, Kenneth
TITLE OF INVENTION: ftsz
FILE REFERENCE: GM10068
CURRENT APPLICATION NUMBER: US/09/120,426
CURRENT FILING DATE: 1998-07-22
EARLIER APPLICATION NUMBER: 60/055,720
EARLIER FILING DATE: 1997-08-12
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 419
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-120-426-2

Query Match 53.3%; Score 472; DB 4; Length 419;
Best Local Similarity 52.6%; Pred. No. 2.6e-46;

Matches 91; Conservative 35; Mismatches 47; Indels 0; Gaps 0;
QY 6 GVEFIYANTDCCALGRSLAPHKITLTKGDKITGKAGSKPELGKRSADQKVIDIOMLQDS 65
DB 38 GVEFIYANTDCCALGRSLAPHKITLTKGDKITGKAGSKPELGKRSADQKVIDIOMLQDS 97
QY 66 NMLFTGGMGGCTGGAAPVAVASVARELGITLVGVSTPFRSGPNRTRLANAGVELAK 125
DB 98 DMVFITAGMGSGGTGAAPVAVARIKADLGLTGVVTRPFGFSGKRGQFAVEGINQJRE 157
QY 126 YVDLIVVNNLALADKSTMLFAFRYADVILEGVKGVTDLIYRPGILNL 178
DB 158 HVDLIIISNNLLEIVDKITLLEALSDADNVLROGVGIDTLITNPGILNL 210

RESULT 4
US-08-883-515-4
Sequence 4, Application US/08883515
Patent No. 5981836
GENERAL INFORMATION:
APPLICANT: Osteryoung, Katherine W
TITLE OF INVENTION: PLANT CHLOROPLAST DIVISION GENES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,515
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 920905.90016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-883-515-4

Query Match 52.3%; Score 463; DB 2; Length 437;
Best Local Similarity 49.7%; Pred. No. 3e-45;
Matches 89; Conservative 36; Mismatches 52; Indels 2; Gaps 1;
QY 2 SLEGEFIVANTDCCALGRS--LAPHKITLTKGDKITGKAGSKPELGKRSADQKVIDIOM 59
DB 99 SEMSGVEFIVNTDQAMRMSVPLPDNRLOKKEITLKGAGAGNPEIGMNAARESEKVEIE 158
QY 60 RMLDSSNMLFTGGMGGCTGGAAPVAVASVARELGITLVGVSTPFRSGPNRTRLANAG 119
DB 159 EALYGSDMVFITAGMGSGGTGAAPVAVARIKADLGLTGVVTRPFGFSGKRGQFAVE 218
QY 120 VKELAKYVDLIVVNNLALADKSTMLFAFRYADVILEGVKGVTDLIYRPGILNL 178
DB 219 LASLDNVDLIVVNNLALADKSTMLFAFRYADVILEGVKGVTDLIYRPGILNL 277

RESULT 5
US-08-883-515-2

Sequence 2, Application US/08883515
Patent No. 5981836
GENERAL INFORMATION:
APPLICANT: Osteryoung, Katherine W
TITLE OF INVENTION: PLANT CHLOROPLAST DIVISION GENES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Charles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,515
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 920905.90016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-883-515-2

Query Match 48.5%; Score 429; DB 2; Length 433;
Best Local Similarity 48.3%; Pred. No. 2.6e-41;
Matches 86; Conservative 40; Mismatches 52; Indels 0; Gaps 0;

QY 1 ASOLEGVEFIANTDCALGRSLAPKHTLTKGKLGAGSKPELGRKSADQKVDIOR 60
DB 94 SSGQSDFVFAINTDSQALQFSAENPLQIGELTGRIGTGNPLLEQAAEESKDAIAN 153
QY 61 MLDQSNMIFITGCGGCTGAAPVAVARELGILTYCVSTPRSEGNPTRLANAGV 120
DB 154 ALKGSDFVFAINTDSQALQFSAENPLQIGELTGRIGTGNPLLEQAAEESKDAIAN 213
QY 121 KELAKYVDLIVPQNQLALADKSTMLAEAFRYADVLEGGVGTDLIVRPGILN 178
DB 214 EKLOKNVDLIVIRNDRLLADDEPTPLQDAFLADVDVLRQGVGISDITITPGLNV 271

RESULT 6
US-09-120-426-4
Sequence 4, Application US/09120426
Patent No. 6197300
GENERAL INFORMATION:
APPLICANT: Fueyo, Joanna Lynn
APPLICANT: Lonetto, Michael A.
APPLICANT: Pearce, Kenneth
TITLE OF INVENTION: fts2
FILE REFERENCE: GM10068
CURRENT APPLICATION NUMBER: US/09/120,426
CURRENT FILING DATE: 1998-07-22
EARLIER APPLICATION NUMBER: 60/055,720
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 223
TYPE: PRT

ORGANISM: Streptococcus pneumoniae
US-09-120-426-4

Query Match 30.6%; Score 271; DB 4; Length 223;
Best Local Similarity 53.6%; Pred. No. 1.9e-23;
Matches 52; Conservative 20; Mismatches 25; Indels 0; Gaps 0;

QY 82 AAPVAVARELGILTYCVSTPRSEGNPTRLANAGKELAKYVDLIVPQNQLAL 141
DB 1 SAPYIARIADLGLTYGVYTRPFGESKRQGFRAVEGIDQREHVDTLLIISNNLEI 60

QY 142 ADKSTMLAEAFRYADVLEGGVGTDLIVRPGILN 178
DB 61 VDKTPLLALSEADNVLRQGVGCTITDLITNPGLINL 97

RESULT 7
US-08-961-083-124
Sequence 124, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 124:
SEQUENCE CHARACTERISTICS:
LENGTH: 283 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-124

Query Match 22.1%; Score 196; DB 4; Length 283;
Best Local Similarity 51.4%; Pred. No. 1.3e-14;
Matches 38; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

QY 105 FRSEGNPTRLANAGVELAKYVDLIVPQNQLALADKSTMLAEAFRYADVLEGVK 164
DB 1 FGEGSKRGQFAVEGIDQREHVDTLLIISNNLEIYDKTPLLALSEADNVLRQGV 60

QY 165 GATDLIVRPGILN 178
DB 61 GITDLITNPGLINL 74

RESULT 8

US-09-308-375-2
; Sequence 2, Application US/09308375
; Patent No. 6300117
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Proteases from Gram-Positive Organisms
; FILE REFERENCE: GC394-PC7
; CURRENT APPLICATION NUMBER: US/09/308,375
; CURRENT FILING DATE: 1999-05-14
; EARLIER APPLICATION NUMBER: EP9719636.4
; EARLIER FILING DATE: 1997-09-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2285
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-308-375-2

Query Match 10.2%; Score 90.5; DB 4; Length 2285;
Best Local Similarity 21.4%; Pred. No. 0.47;
Matches 42; Conservative 32; Mismatches 79; Indels 43; Gaps 7;

OY 1 ASOL-EGVEFIYANTDQALGRSLAPHKI-----TLGKDIK----- 36
DB 2091 AKQLEFSKEI--NIMMESIGKISINMLIDKLKESNALNTAVKNTGKVKVSSPASGGY 2148
OY 37 ---GLAGSKPELGKRSADQKVDIORMLQDSNMLFITGGMGGCTGGAAPVAVAREL 93
DB 2149 TGTGAGAGLAFHDKELINKDITANILDTYKAV-----RETAVDSPKMGGVKLA 2201
OY 94 GILTVGVSTPFRESEGNRTLRANAGYKELAKYVDTLIVPNOMLALADKSTMLEAFR 153
DB 2202 DIKKGITSPISLIVPNVMSLNTSLNPLK-----LEIPSKTIASGDKTINLTNTHF 2256
OY 154 VADDVLEGGVGTDL 169
DB 2257 I--DKLIGESGARSM 2270

RESULT 9

US-09-134-001C-4951
; Sequence 4951, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4951
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4951

Query Match 8.9%; Score 79; DB 4; Length 388;
Best Local Similarity 23.5%; Pred. No. 0.75;
Matches 32; Conservative 32; Mismatches 48; Indels 24; Gaps 6;

OY 11 VANTDQALGRSLAPHKITLGDITGGLAGSKPELGKRSADQKVDIORMLQDSNMLFI 70
DB 24 IANTNGOYLSKLFN---SIGKSVETHVIGDNP-----QRLEVIYKQCLSRFDITVL 72
OY 71 TGMGSGCTGGAAPVAVAREL--ILT---GVVSTPFRSGPNRTLRANAGYKELA 124
DB 73 TGGGLP---TKDILTKHTVAKYGLKMLVDEASLNTKYNFKEGGDMT---SNKQQA 125

OY 125 KYVDLIVPNOMLLA 140
DB 126 LVIEDAIVLPKNKGMA 141

RESULT 10

US-09-189-527-11
; Sequence 11, Application US/09189527A
; Patent No. 6387639
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
; FILE REFERENCE: SLK98-01
; CURRENT APPLICATION NUMBER: US/09/189,527A
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 271
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-11

Query Match 8.7%; Score 77; DB 4; Length 271;
Best Local Similarity 29.1%; Pred. No. 0.76;
Matches 30; Conservative 9; Mismatches 30; Indels 34; Gaps 5;

OY 37 GLAGSKPELGKRSADQKVDIORMLQDSNMLFITG---GMGGCTGGA---APVAVS 88
DB 32 GLGVVSPGAGGRVSSHSLHLTR-----LTGPFVGTGNGTCASAPATHEIPEYAS 82
OY 89 VARE-----LGILTVGVSTPFRESEGNRTLRANAGYVE 122
DB 83 ILRECCPSPRGRVGLVGTG-----HGDSQVAVASNGRKE 117

RESULT 11

US-09-134-001C-3034
; Sequence 3034, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3034
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3034

Query Match 8.6%; Score 76.5; DB 4; Length 422;
Best Local Similarity 23.0%; Pred. No. 1.17;
Matches 42; Conservative 25; Mismatches 61; Indels 55; Gaps 8;

OY 34 ITKGLAGSKPELGKRSADQKVDIORMLQDSNMLFITGGMGGCTGGAAPVAVAREL 93
DB 214 ILRGTVASVPPGGGRKHPNDELIOI---DTTNILFIIGAFDG-----IDEVIKRRL 262
OY 94 GILTVGVSTPFRESEGNRTLRANAGYKELAKY-----VDTLIVPNOML 138
DB 263 GERVIGFASNE-ADKYDEALIEQIRPEDLOSGLIPEIGRVPITVAMLETLDVAALKNI 321

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2003, 05:46:58 ; Search time 65 seconds
(without alignments)
364.902 Million cell updates/sec

Title: US-09-770-509-2

Perfect score: 885

Sequence: 1 ASQLEGVEFIVANTDQALG.....LLGKGVTDLIVRGLINL 178

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID2/gcgdata/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	57.6	391	23	ABBA8477
2	495	55.9	410	22	AAU35046
3	493	55.9	411	23	AAU33413
4	493	55.7	416	23	ABBS5218
5	485	54.8	379	22	AAAG8170
6	480	54.2	390	21	AAV91108
7	480	54.2	390	21	AAV44710
8	480	54.2	390	22	AAU34274
9	480	54.2	392	22	AAU37130
10	480	54.2	394	23	ABP39989

11	472	53.3	335	19	AAW80614
12	472	53.3	419	20	AAW99120
13	472	53.3	419	21	AAV44711
14	472	53.3	419	22	AAU37886
15	470	53.1	442	22	AAU32106
16	466	52.7	421	22	AAU60468
17	464	52.4	394	22	AAU56414
18	464	52.4	439	23	ABP25525
19	463	52.3	397	21	AAU801338
20	463	52.3	437	19	AAU41733
21	456.5	51.6	427	23	ABP25524
22	452	51.1	383	21	AAU15908
23	452	51.1	383	21	AAV91110
24	452	51.1	383	22	AAU34440
25	452	51.1	413	21	AAU801330
26	447	50.5	383	23	AAU76932
27	447	50.5	383	23	AAU76923
28	447	50.5	383	23	AAU76924
29	447	50.5	383	23	AAU76925
30	447	50.5	421	22	AAU76935
31	441	49.8	374	19	AAW70503
32	441	49.8	374	19	AAW41787
33	440	49.7	361	19	AAW51423
34	439	49.6	325	21	AAU801332
35	435	49.2	363	21	AAU46656
36	435	49.2	372	21	AAU46655
37	435	49.2	433	21	AAU46654
38	435	49.2	433	21	AAU801337
39	433	48.9	411	21	AAU801329
40	429	48.5	433	19	AAW41732
41	424.5	48.0	395	22	AAU86802
42	379	42.8	385	22	AAU35960
43	323.5	36.6	413	22	AAU86344
44	304	34.6	357	21	AAU801333
45	285	32.2	267	21	AAU804705

ALIGNMENTS

RESULT 1
ABBA8477
ID ABBA8477 standard; Protein: 391 AA.
XX
AC ABBA8477;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #1181.
XX
XX Antibacterial; gene therapy: vaccine: biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN W0200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001MO-FR01118.
XX
PR 11-APR-2000; 2000FR-0004629.
XX
PA (INSP) INST PASTEUR.
XX
PI Buchliesser C, Frangoul L, Couve E, Rusnok C, Eshti H, Dehoux P;
PI Dussurget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz T, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Madueno E, De Pablo B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;

S.pneumoniae filam
Streptococcus pneu
Monomeric filament
Streptococcus pneu
C glutamicum prote
Propionibacterium
Pseudomonas aerugi
Streptococcus poly
fcs22 polypeptide
Arabidopsis chloro
Streptococcus poly
E. coli proliferat
Escherichia coli m
E. coli cellular p
xanfcs2 polypepti
wild-type E. coli
E. coli ftsz prote
E. coli ftsz prote
Hemophilus influe
Brevibacterium fla
Neisseria meningit
fts21 polypeptide
Arabidopsis thalia
Arabidopsis thalia
fts21 polypeptide
Bartsz1 polypeptid
Arabidopsis chloro
Putative P. abyssi
Helicobacter pylori
Putative P. abyssi
fts21 polypeptide
Arabidopsis thalia

XX WPI; 2002-010914/01.
 XX Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment
 XX PT and prevention of *Listeria* and related bacterial infections, and
 XX PT related polypeptides -
 XX Claim 6; SEQ ID No 1182; 192pp; French.
 XX
 XX The present invention relates to the genome sequence of *Listeria*
 CC *monocytogenes* EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in *L.*
 CC *monocytogenes* and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies. Identification of *L. monocytogenes* and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate *L. monocytogenes*-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by *L.*
 CC *monocytogenes* and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 SQ Sequence 391 AA;
 Query Match 57.6%; Score 510; DB 23; Length 391;
 Best Local Similarity 56.0%; Pred. No. 1.5e-45;
 Matches 98; Conservative 35; Mismatches 42; Indels 0; Gaps 0;
 QY 4 LGEVEFIYVNTDQALGRSLAPHKITLTKGDKGAGSKPELGRSAEQKVDIORMQ 63
 DB 35 VGVGEFISVNTDQALNTAKAEKTKLOICTKTRGLGAGVPEIGKKAEESEQVFEALK 94
 QY 64 DSNMFLITGGMGSGCTGGAAPVAVASVARELGILTVGVVSPFSEGPNTRLANAGVKEL 123
 DB 95 GSDMVFVTAGMGSGTGTGAAPVIAOLIAKEKMGALTVGVVTRFEGEPKTRKQALGTTEAM 154
 QY 124 AKYVDLLIYVNPQNLALADKSTMTLEAFRYADVLLLEGVKGVTDLIVRGLINL 178
 DB 155 KEAVDTLLIYVNPDLRLQIVDKNTMLEAFREADVLLRQGVGQISDLTAVPGLINL 209
 RESULT 2
 ID AU035046 standard; Protein; 410 AA.
 AC AU035046;
 XX
 DT 13-FEB-2002 (first entry)
 DE Enterococcus faecalis cellular proliferation protein #333.
 KM Antisense; prokaryotic cellular proliferation protein;
 KM antibiotic; antibacterial; drug design.
 OS Enterococcus faecalis.
 XX
 PN WO200170955-A2.
 PD 27-SEP-2001.
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 XX
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 PI
 DR WPI; 2001-611495/70.
 DR N-PSDB: AAS52905.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 10639; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences.
 CC
 SQ Sequence 410 AA;
 Query Match 55.9%; Score 495; DB 22; Length 410;
 Best Local Similarity 57.1%; Pred. No. 6.2e-44;
 Matches 100; Conservative 28; Mismatches 47; Indels 0; Gaps 0;
 QY 4 LGEVEFIYVNTDQALGRSLAPHKITLTKGDKGAGSKPELGRSAEQKVDIORMQ 63
 DB 36 VKGEFETANTDVQALKHSKAEYVQLGPKYTRGAGSQPVGOKAAESESQVISESIQ 95
 QY 64 DSNMFLITGGMGSGCTGGAAPVAVASVARELGILTVGVVSPFSEGPNTRLANAGVKEL 123
 DB 96 GADMIFITAGMGSGTGTGAAPVAVAKIAKELGALTGGVTRFSPFEGPKRGFRPAESGIALL 155
 QY 124 AKYVDLLIYVNPQNLALADKSTMTLEAFRYADVLLLEGVKGVTDLIVRGLINL 178
 DB 156 KENVDTLLIISNNRLLEVVDKTPMLEAFREADVLLRQGVGQISDLTAPGVYNL 210
 RESULT 3
 ID AU033413 standard; Protein; 411 AA.
 AC AU033413;
 XX
 DT 14-FEB-2002 (first entry)
 DE Enterococcus faecalis cellular proliferation protein #49.
 KM Antisense; prokaryotic cellular proliferation protein;
 KM antibiotic; antibacterial; drug design.
 OS Enterococcus faecalis.
 XX
 PN WO200170955-A2.
 PD 27-SEP-2001.

XX 21-MAR-2001; 2001MO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 PI Haselbeck R, Ohlsen KL, Zyckind JW, Wall D, Trawick JD, Carr GJ,
 PI Yamamoto RT, Xu HH.
 XX
 DR WPI: 2001-611495/70.
 DR N-PSDB; AAS51272.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 4909; 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences.

SO Sequence 411 AA:

Query Match 55.9%; Score 495; DB 22; Length 411;
 Best Local Similarity 57.1%; Pred. No. 6.3e-44;
 Matches 100; Conservative 28; Mismatches 47; Indels 0; Gaps 0;

OY 4 LGEVEFIVANTDQALGRSLAPKHTLTKDITKGLGAGSKPELGKRSADQKVDIORMIQ 63
 DB 36 VSGVEFIVANTDQALGRSLAPKHTLTKDITKGLGAGSKPELGKRSADQKVDIORMIQ 95
 OY 64 DSNMFLITGGMGGCTGTGAAPVAVASAREIGLTVGVSPFSEGNRRLANAGVKEL 123
 DB 96 GADMIFITGGMGGCTGTGAAPVAVAKIAKELGALTGVVTRPFSEGNRRLANAGVKEL 155
 OY 124 AKYVDLIVVPNONLALADKSTTMEAFRYADVDLLEGYKGVYDILVRPGLINL 178
 DB 156 KENVDTLLIISNNRLLEVDKTPMLEAFREADNVLRQGVQGISDILITAFGYVNL 210

RESULT 4
 ABB55218
 ID ABB55218 standard; Protein; 416 AA.
 XX
 AC ABB55218;
 XX
 DT 16-MAY-2002 (first entry)
 XX
 DE Lactococcus lactis protein ftsz.
 XX

KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 XX
 OS Lactococcus lactis IL1403.
 XX
 PN FR807446-A1.
 XX
 PD 12-OCT-2001.
 XX
 PF 11-APR-2000; 2000FR-0004630.
 XX
 PR 11-APR-2000; 2000FR-0004630.
 XX
 PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 PI WPI: 2002-043418/06.
 XX
 DR New nucleotide sequence useful in the identification or Lactococcus
 DR lactis and related species -
 XX
 PT Claim 6; SEQ ID No 1920; 2504pp; French.

CC The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify Lactococcus lactis or
 CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC WO200177334 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

SO Sequence 416 AA:

Query Match 55.7%; Score 493; DB 23; Length 416;
 Best Local Similarity 56.0%; Pred. No. 1e-43;
 Matches 98; Conservative 29; Mismatches 48; Indels 0; Gaps 0;

OY 4 LGEVEFIVANTDQALGRSLAPKHTLTKDITKGLGAGSKPELGKRSADQKVDIORMIQ 63
 DB 36 VSGVEFIVANTDQALGRSLAPKHTLTKDITKGLGAGSKPELGKRSADQKVDIORMIQ 95
 OY 64 DSNMFLITGGMGGCTGTGAAPVAVASAREIGLTVGVSPFSEGNRRLANAGVKEL 123
 DB 96 GADMIFITGGMGGCTGTGAAPVAVAKIAKELGALTGVVTRPFSEGNRRLANAGVKEL 155
 OY 124 AKYVDLIVVPNONLALADKSTTMEAFRYADVDLLEGYKGVYDILVRPGLINL 178
 DB 156 RANVDLILLIISNNRLLEVDKTPMLEAFREADNVLRQGVQGYTDILITNGMINL 210

RESULT 5
 AAG81170
 ID AAG81170 standard; Protein; 379 AA.
 XX
 AC AAG81170;
 XX
 DT 04-SEP-2001 (first entry)
 XX
 DE Mycobacterium tuberculosis potential drug target protein SPO ID 221.
 XX
 KW Drug target; growth; organism viability; characterisation.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN WO200135317-A1.
 XX
 PD 17-MAY-2001.
 XX
 DE 13-NOV-2000; 2000MO-US31152.
 XX
 PF

XX 12-NOV-1999; 99US-0165086.
 PR 12-NOV-1999; 99US-0165124.
 PR 01-FEB-2000; 2000US-0179531.
 XX
 XX (REGC) UNIV CALIFORNIA.
 XX
 PI Eisenberg D, Rotstein SH, Marcotte EM;
 XX
 DR WPI: 2001-329193/34.
 DR N-PSDB: AAH52021.
 XX
 PT Identifying nucleotide or polypeptide sequence for use as drug target,
 PT involves providing algorithm that analyzes a functional relationship
 PT between nucleotide or polypeptide sequences, and comparing the
 PT sequences
 PS Disclosure; Page 172; 207pp; English.
 XX
 XX This invention relates to a method for identifying a nucleotide or
 CC polypeptide sequence that may be a drug target, or essential for growth
 CC or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092
 CC represent DNA encoding proteins AAG81096 AAG81241, Mycobacterium
 CC tuberculosis proteins which are potential drug targets. The DNA and
 CC protein sequences are used to illustrate the method of the invention. The
 CC method involves providing an unknown nucleotide or polypeptide sequences,
 CC and comparing it to a number of sequences along with at least one
 CC algorithm capable of analyzing a functional relationship between
 CC nucleotide and polypeptide sequences. The method is useful for
 CC characterizing the function of nucleic acids and polypeptides that may be
 CC useful as a target for a drug or essential for the growth or viability of
 CC an organism.
 XX
 XX Sequence 379 AA:
 SQ
 Query Match 54.8%; Score 485; DB 22; Length 379;
 Best Local Similarity 54.3%; Pred. No. 6.3e-43;
 Matches 95; Conservative 35; Mismatches 45; Indels 0; Gaps 0;
 OY 4 LEGVEFVANTDQALGRSLAPKHTLTKDITKLGAGSKPELGRSAEQKVDIORMQ 63
 DB 32 LKGVETFIATNTDQALMLMSADYKLDVGRSTRGLGAGDEYVGRKAEDAKDEIBLLR 91
 OY 64 DSNMLFTGGGSGGTCTGAAPVAVASVARELGILTVGVSTPSPRSEGNPRTLANAGVKEL 123
 DB 92 GADMFVETSGEGGTGTGAPVAVASIAKRGALTGVGTVTRFSPGKRRSMQAEINGIAL 151
 OY 124 AKYVDTLIVPNQNLALADKSTMLEAFRYADDVLEGVKGVDTLIVRPGINL 178
 DB 152 RSCDTLIVIPNDRLIQMGDAVSLMDAFRSADDEVLLNGVGITDLITPGLINV 206
 RESULT 6
 AAY91108
 ID AAY91108 standard; Protein; 390 AA.
 XX
 AC AAY91108;
 XX
 DT 15-SEP-2000 (first entry)
 XX
 DE Staphylococcus aureus monomeric FtsZ protein sequence SEQ ID NO:2.
 XX
 XX Staphylococcus aureus; monomeric FtsZ; monomeric FtsA; antibacterial;
 KW Escherichia coli; multimeric protein; bacterial infection; vaccine;
 XX microbial infection; screening.
 XX
 OS Staphylococcus aureus.
 XX
 XX WO200024252-A1.
 PN
 XX 04-MAY-2000.
 PD
 XX 20-OCT-1999; 99WO-US24653.
 PF

XX 23-OCT-1998; 98US-0105315.
 PR 19-APR-1999; 99US-0129965.
 PR 03-MAY-1999; 99US-0132333.
 XX
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Van K, Pearce KH;
 XX
 DR WPI: 2000-350517/30.
 DR N-PSDB: AAA39395.
 XX
 PT Multimeric FtsZ:FtsA polypeptide, useful as a vaccine for treating
 PT bacterial infections and for diagnosing diseases associated with
 PT bacterial infection
 PS Claim 1; Page 2; 45pp; English.
 XX
 XX The present invention describes isolated multimeric FtsZ:FtsA
 CC polypeptides (I) comprising the protein sequences given in AAY91108,
 CC AAY91109, AAY91110 and AAY91111. Also described are: (1) an antibody
 CC (Ab) specific for (I); (2) a method (M1) for treating an individual
 CC requiring enhanced or reduced activity or expression of (I) comprising
 CC administering an agonist or antagonist to (I), respectively; (3) a
 CC method (M2) for diagnosing a disease or susceptibility to a disease
 CC related to expression or activity of (I) comprising determining the
 CC presence or absence of a mutation in the nucleotide sequence encoding
 CC (I) and/or analyzing for the presence or quantity of (I) in a sample;
 CC and (4) a method (M3) for screening to identify compounds that activate
 CC or inhibit the function of (I) selected from: (a) measuring the binding
 CC of a candidate compound to the polypeptide using a label; (b) measuring
 CC the binding of a candidate compound to (I) in the presence of a labeled
 CC competitor; (c) testing whether the candidate compound results in a
 CC signal generated by activation or inhibition of (I); or (d) detecting
 CC the effect of a candidate compound on the production of mRNA encoding
 CC (I) using e.g. enzyme linked immunosorbent assay (ELISA). AAY91108 to
 CC AAY91111 are useful as vaccines for treating microbial infections such
 CC as Staphylococcus, Streptococcus and Escherichia coli infections. The
 CC polypeptides are also useful for screening for antibacterial compounds
 CC and for detecting diseases associated with microbial infections. The
 CC present sequence represents Staphylococcus monomeric FtsZ.
 XX
 XX Sequence 390 AA:
 SQ
 Query Match 54.2%; Score 480; DB 21; Length 390;
 Best Local Similarity 54.3%; Pred. No. 2.3e-42;
 Matches 95; Conservative 33; Mismatches 47; Indels 0; Gaps 0;
 OY 4 LEGVEFVANTDQALGRSLAPKHTLTKDITKLGAGSKPELGRSAEQKVDIORMQ 63
 DB 35 MNVEFETIATNTDQALMLMSADYKLDVGRSTRGLGAGANPGRKAEDAKDEIBLLR 94
 OY 64 DSNMLFTGGGSGGTCTGAAPVAVASVARELGILTVGVSTPSPRSEGNPRTLANAGVKEL 123
 DB 95 GADMFVETSGEGGTGTGAPVAVAKIAKEMGALTGVGTVTRFSPGKRRSQTAAGVEAM 154
 OY 124 AKYVDTLIVPNQNLALADKSTMLEAFRYADDVLEGVKGVDTLIVRPGINL 178
 DB 155 KAAVDTLIVIPNDRLIDVSKSTPMMEAFKADNVLLRGVQGISDILAVSGEYNL 209
 RESULT 7
 AAY44710
 ID AAY44710 standard; Protein; 390 AA.
 XX
 AC AAY44710;
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE S. pneumoniae monomeric filamentation temperature sensitive protein.
 XX
 XX Monomeric filamentation temperature sensitive protein;
 KW FtsZ; dental treatment; vaccine; wound infection; gastric ulcer;
 KM

DE S. pneumoniae filamentous temperature sensitive, cell septation protein.
XX Streptococcus pneumoniae protein: recombinant; gene expression;
KM DNA chip; virulence; antibody; infection; detection; treatment.
XX
OS Streptococcus pneumoniae.
XX
PN WO9826072-A1.
XX
PD 18-JUN-1998.
XX
PF 09-DEC-1997; 97MO-US22578.
XX
PR 13-DEC-1996; 96US-0036281.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Balazs RH, Burgett SG, Dehoff BS, Hoskins JA, Jaskunas SF;
PI Mills BJ, Norris FH, Peery RB, Rostock PR;
PI Skatrud PL, Smith MC, Solenberg RJ, Treadway RJ;
PI Young Bellido ML;
XX
DR WPI: 1998-348529/30.
DR N-PSDB; AAV655296.
XX
XX Streptococcus pneumoniae nucleic acid sequences - used in DNA chips
PT for evaluating gene expression, and identification of virulence
PT genes
PS Claim 3; Pages 182-183; 333pp; English.
XX
XX This sequence represents a S. pneumoniae filamentous temperature
CC sensitive, cell septation protein. The invention provides DNA sequences
CC (AAV655201 to AAV65304) from the S. pneumoniae genome and corresponding
CC protein sequences (AAV60605 to AAV60728). A recombinant host containing a
CC vector comprising any of the above nucleic acids can be used for the
CC recombinant expression of the protein sequences. The invention also
CC provides a DNA chip having arrayed on it at least 15 base pair fragment
CC of any one or more of these DNA sequences. The DNA chip can be used
CC methods for evaluating gene expression in S. pneumoniae and for
CC identifying virulence genes in S. pneumoniae. Antibodies that
CC selectively bind to the above proteins or peptide fragments can be used
CC to treat S. pneumoniae infection. The antibodies can also be used to
CC detect S. pneumoniae cells.
XX
XX Sequence 335 AA:
SQ
Query Match 53.3%; Score 472; DB 19; Length 335;
Best Local Similarity 52.6%; Pred. No. 1.3e-41;
Matches 91; Conservative 35; Mismatches 47; Indels 0; Gaps 0;
QY 6 GVEFIYANTDCAQLGSLAPHKITLTKGDKGAGSKPELGRSAQKVDIORMQDS 65
DB 38 GVEFIYANTDVQALSTKAEIVIQGLPKILRGAGGQPEVGRKADESEETLEAISGA 97
QY 66 NMLFTTGGGGGCTCGAPVAVASVARELITLVGVVSTPPRSGPNRTRLANAGVKELAK 125
DB 98 DMVFTTAGGGGSGTGAAPVIARIADLGLTGVVTRPFGFSGSKRGQFAVEGINQLORE 157
QY 126 YVDTLIVPNOMLALADKSTMLAEFRYADVLLBESGVDTLIVRGLINL 178
DB 158 HVDTLITISNNLLEIVDKTLPLEALSEADNVLRQGVGQITDLITNPGILNL 210

RESULT 12

AAW99120
ID AAW99120 standard; Protein; 419 AA.
XX
XX AAW99120;
AC
XX
DT 17-MAY-1999 (first entry)
XX
DE Streptococcus pneumoniae ftsz protein.

XX Streptococcus pneumoniae; ftsz; antibacterial; microbial disease;
KM otitis media; bacteremia; conjunctivitis; pneumonia; sinusitis;
KM pleural empyema; endocarditis; meningitis.
XX
OS Streptococcus pneumoniae.
XX
PN EP899334-A2.
XX
XX 03-MAR-1999.
XX
PD 30-JUL-1998; 98BP-0306077.
XX
PF 12-AUG-1997; 97US-0055720.
XX
PR (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Fueyo JL, Lonetto MA;
PI WPI: 1999-144805/13.
DR N-PSDB; AAX19269.
XX
XX New Streptococcus pneumoniae ftsz polypeptide and polynucleotide -
PT useful as diagnostic reagents and for prevention and treatment of
PT Streptococcal infections which cause bacteraemia, otitis media and
PT meningitis
PS Claim 1; Page 4; 37pp; English.
XX
XX The present sequence is ftsz polypeptide isolated from Streptococcus
CC pneumoniae. ftsz polynucleotides and polypeptides are useful for
CC diagnosing susceptibility to diseases by detecting mutations or
CC polymorphisms in the ftsz gene or analysing for the presence of amount of
CC ftsz polypeptide expressed in a patient sample. ftsz PCR probes are
CC useful for diagnosing diseases, and can characterise the response of
CC the infectious organism to drugs. ftsz polypeptides and polynucleotides
CC are also useful for screening for antagonists, agonists and drugs against
CC infectious micro-organisms. ftsz agonists and antagonists are
CC bacteriostatic and bacteriocidal compounds which can be used in treatment
CC to enhance (agonist) or block (antagonist or antisense sequence) ftsz
CC activity, therefore treating microbial diseases, especially Streptococcus
CC pneumoniae diseases including otitis media, bacteraemia, conjunctivitis,
CC pneumonia, sinusitis, pleural empyema, endocarditis and especially
CC meningitis. Epitopes of ftsz polypeptides and polynucleotides are useful
CC immunogens for producing anti-ftsZ antibodies for prevention of bacterial
CC infections, and ftsz polynucleotides can be used in genetic immunisation
CC (gene therapy) to prevent infections. ftsz polypeptides, polynucleotides
CC and their (ant)agonists can prevent adhesion of bacteria to matrix
CC proteins, and are useful for use on wounds and body implants to prevent
CC bacterial infection. ftsz polypeptides and polynucleotides may also be
CC used as reagents for differential screening methods e.g. using ftsz
CC probes in RT-PCR to identify and quantify genes expressed in bacterial
CC tissue.
XX
XX Sequence 419 AA:
SQ
Query Match 53.3%; Score 472; DB 20; Length 419;
Best Local Similarity 52.6%; Pred. No. 1.8e-41;
Matches 91; Conservative 35; Mismatches 47; Indels 0; Gaps 0;
QY 6 GVEFIYANTDCAQLGSLAPHKITLTKGDKGAGSKPELGRSAQKVDIORMQDS 65
DB 38 GVEFIYANTDVQALSTKAEIVIQGLPKILRGAGGQPEVGRKADESEETLEAISGA 97
QY 66 NMLFTTGGGGGCTCGAPVAVASVARELITLVGVVSTPPRSGPNRTRLANAGVKELAK 125
DB 98 DMVFTTAGGGGSGTGAAPVIARIADLGLTGVVTRPFGFSGSKRGQFAVEGINQLORE 157
QY 126 YVDTLIVPNOMLALADKSTMLAEFRYADVLLBESGVDTLIVRGLINL 178
DB 158 HVDTLITISNNLLEIVDKTLPLEALSEADNVLRQGVGQITDLITNPGILNL 210

```

RESULT 13
AAV44711
ID AAV44711 standard; Protein; 419 AA.
XX
AC AAV44711;
XX
DT 25-APR-2000 (first entry)
XX
DE Monomeric filamentation temperature sensitive protein.
XX
KW Monomeric filamentation temperature sensitive protein;
XX FtsZ; dental treatment; vaccine; wound infection; gastric ulcer;
XX gastrointestinal cancer; Helicobacter pylori infection; gastritis;
XX antibacterial; antiulcer; anti-inflammatory.
XX
OS Staphylococcus aureus.
XX
PN WO200001801-A1.
XX
PD 13-JAN-2000.
XX
PF 30-JUN-1999; 99WO-US14890.
XX
PR 02-JUL-1998; 98US-0091680.
XX
PA (SMIR) SMITHKLINE BEECHAM CORP.
XX
PI Hensley CP, Pearce KH, Sossong TM;
XX
DR WPI: 2000-171008/15.
XX N-PSDB; AAV49959.
XX
PT Isolated polypeptide useful as a vaccine and for screening
XX antimicrobial compounds for treating microbial diseases -
XX
PS Claim 1; Page 4; 52pp; English.
XX
CC The present sequence is the monomeric filamentation
CC temperature sensitive protein (FtsZ). This is related by amino acid
CC sequence homology to Lactococcus lactis FtsZ AAV74322. Multimeric FtsZ
CC protein is useful in identifying compounds that modulate its function.
CC The protein and its variants are useful in diagnosing and treating
CC microbial diseases and for prophylactic use in dental treatment. They
CC can be used as vaccines to prevent wound infections by bacteria
CC especially S. pneumoniae. Agonists and antagonists are useful in
CC treating diseases caused by Helicobacter pylori infections e.g.
CC gastrointestinal cancer, gastric ulcer and gastritis.
CC Note: The specification is unclear as to whether this sequence was
CC isolated from Streptococcus pneumoniae or Staphylococcus aureus.
XX
SQ Sequence 419 AA;
XX
Query Match 53.3%; Score 472; DB 21; Length 419;
Best Local Similarity 52.6%; Pred. No. 1.8e-41;
Matches 91; Conservative 35; Mismatches 47; Indels 0; Gaps 0;
XX
QY 6 GVEFTVANTDQALGRSLAPHKITLKGDTKGLGAGSKPELGKRSADQKVDIRMLDGS 65
DB 38 GVEFTIANDVQALSTKAEVYIOLGPKLIRGLGAGGQPEVGRKAAESEEFTLTERISA 97
QY 66 NMLFITGGMGGGCTGAAPVAVASARELGITLVGVSTPRSEGNRRLANAGYKELAK 125
DB 98 DMVFTTAGMGSGGCTGAAPVARIADKDALTVGVTRPFEFGSKRGQFAVEGINQLRE 157
QY 126 YVDTLIVPNOMLIALADKSTTMLEAFRYADVDVLEBGKGVTDLIVRGLINL 178
DB 158 HVDTLIIISNNLILETVDKTPLLLEALSEADNVLRQGVGTTDLITNGLINL 210
XX
RESULT 14
AAU37886
ID AAU37886 standard; Protein; 419 AA.
XX

```

```

AC AAU37886;
XX
DT 14-FEB-2002 (first entry)
XX
DE Streptococcus pneumoniae cellular proliferation protein #215.
XX
KW Antisense; prokaryotic cellular proliferation protein;
XX antibiotic; antibacterial; drug design.
XX
OS Streptococcus pneumoniae.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
XX
PR 23-MAY-2000; 2000US-206848P.
XX
PR 26-MAY-2000; 2000US-207727P.
XX
PR 23-OCT-2000; 2000US-242578P.
XX
PR 27-NOV-2000; 2000US-253625P.
XX
PR 22-DEC-2000; 2000US-257931P.
XX
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX
DR WPI: 2001-611495/70.
XX N-PSDB; AAS55745.
XX
PT New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 13479; 51pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antipodids capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 419 AA;
XX
Query Match 53.3%; Score 472; DB 22; Length 419;
Best Local Similarity 52.6%; Pred. No. 1.8e-41;
Matches 91; Conservative 35; Mismatches 47; Indels 0; Gaps 0;
XX
QY 6 GVEFTVANTDQALGRSLAPHKITLKGDTKGLGAGSKPELGKRSADQKVDIRMLDGS 65
DB 38 GVEFTIANDVQALSTKAEVYIOLGPKLIRGLGAGGQPEVGRKAAESEEFTLTERISA 97
QY 66 NMLFITGGMGGGCTGAAPVAVASARELGITLVGVSTPRSEGNRRLANAGYKELAK 125
DB 98 DMVFTTAGMGSGGCTGAAPVARIADKDALTVGVTRPFEFGSKRGQFAVEGINQLRE 157
QY 126 YVDTLIVPNOMLIALADKSTTMLEAFRYADVDVLEBGKGVTDLIVRGLINL 178
XX

```

Db 158 HVDVLLIISNNNLEIYDKKTPLEALSEADNVLRGQVGIIDLTNPGLINL 210

Db 152 KEVCDTILIVIPNDRLLELGDANLSIMEAFRADEVLHNGVGIITLITPGVINY 206

Search completed: June 2, 2003, 08:33:21
Job time : 72 secs

RESULT 15

AAG92106 ID AAG92106 standard; Protein: 442 AA.

AC AAG92106;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 5860.

KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

OS Corynebacterium glutamicum.

PN EP108790-A2.

PD 20-JUN-2001.

PF 18-DEC-2000; 2000EP-0127688.

PR 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.

PR 03-AUG-2000; 2000JP-0280988.

PA (KYOW) KYOWA HAKKO KOGYO KK.

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI: 2001-376931/40.

DR N-PSDB; AAH67325.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying

PT mutation point of a gene, measuring expression of a gene, analysing

PT expression profile or pattern of a gene and identifying homologous gene

PT Claim 17; SEQ ID NO: 5860; 246bp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein

CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These

CC are useful for identifying the mutation point of a gene derived from a

CC mutant of coryneform bacterium, measuring expression amount and

CC analysing the expression profile or expression pattern of a gene derived

CC from Coryneform bacterium, and identifying a homologue of a gene derived

CC from Coryneform bacterium. Coryneform bacteria are useful for producing

CC amino acids, nucleic acids, vitamins, saccharides and organic acids,

CC particularly L-lysine. The present sequence is a protein described

CC in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from the

CC European Patent Office.

SO Sequence 442 AA;

Query Match 53.18; Score 470; DB 22; Length 442;

Best Local Similarity 52.68; Pred. No. 3,1e-41;

Matches 92; Conservative 36; Mismatches 47; Indels 0; Gaps 0;

4 LECGEFVANTDOALGRSLAPHKITLGDITKLGAGSKRPELGRSAEQKVDIORMIQ 63

32 LKGEVEFLAVNTDSOALFSDADVRLDGLRENTRLGAGANPEVGRASAEHKNIEIERIK 91

64 DSNMLFITGGGGTCTGAPVAVSVARELGIITGVVSTPFRSEGPNTRLANGVKEI 123

92 GADMVETVINGEGGTGTGAPVAVGIAKKKALITIGVTKRFEFGRRRTROAEIGIAAL 151

124 AKYVDLIVIPNDRLLELGDANLSIMEAFRADEVLHNGVGIITLITPGVINY 178

;; TITLE OF INVENTION: OOMYCETE FT5Z-MT AS A TARGET FOR
;; FILE REFERENCE: NADII.018A
;; CURRENT APPLICATION NUMBER: US/09/770.509
;; CURRENT FILING DATE: 2001-01-25
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 17
;; LENGTH: 401
;; TYPE: PRF
;; ORGANISM: Mallomonas splendens
US-09-770-509-17

Query Match 59.5%; Score 527; DB 9; Length 401;
Best Local Similarity 58.4%; Pred. No. 4.6e-47;
Matches 104; Conservative 27; Mismatches 47; Indels 0; Gaps 0;

QY 1 ASOLEVEFTVAMTDCALGRSLAPHKITITGDKITGAGSKPELGKRSABQKVDIORM 60
DB 100 AKRLSVEFTVAMTDCALGRSLAPHKITITGDKITGAGSKPELGKRSABQKVDIORM 159
QY 61 MLDNSMLFTTGGMGCTGGAAPVAVASVARELGILTVGVSTPFRSEGNPRTRLANAGV 120
DB 160 YIADAMVFTTAGMGCTGGAAPVAVASVARELGILTVGVSTPFRSEGNPRTRLANAGV 219
QY 121 KEIAKYVDLIVPNOMLALADKSTMLAEFRYADVLLBQKGVTDLIVRGLINL 178
DB 220 RSLDEVDLIVPNOMLALADKSTMLAEFRYADVLLBQKGVTDLIVRGLINL 277

RESULT 3
US-09-770-509-14

;; Sequence 14, Application US/09770509
;; Publication No. US20030082657A1
;; GENERAL INFORMATION:
;; APPLICANT: Katagiri, F.
;; TITLE OF INVENTION: OOMYCETE FT5Z-MT AS A TARGET FOR
;; FILE REFERENCE: NADII.018A
;; CURRENT APPLICATION NUMBER: US/09/770.509
;; CURRENT FILING DATE: 2001-01-25
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 14
;; LENGTH: 452
;; TYPE: PRF
;; ORGANISM: Rickettsia prowazekii
US-09-770-509-14

Query Match 58.5%; Score 518; DB 9; Length 452;
Best Local Similarity 55.9%; Pred. No. 4.8e-46;
Matches 99; Conservative 38; Mismatches 40; Indels 0; Gaps 0;

QY 2 SOLEGEFTVAMTDCALGRSLAPHKITITGDKITGAGSKPELGKRSABQKVDIORM 61
DB 36 ANLQAGNEFTVAMTDCALGRSLAPHKITITGDKITGAGSKPELGKRSABQKVDIORM 95
QY 62 LODSNMLFTTGGMGCTGGAAPVAVASVARELGILTVGVSTPFRSEGNPRTRLANAGV 121
DB 96 LENSVMVFTTAGMGCTGGAAPVAVASVARELGILTVGVSTPFRSEGNPRTRLANAGV 155
QY 122 ELAKYVDLIVPNOMLALADKSTMLAEFRYADVLLBQKGVTDLIVRGLINL 178
DB 156 ELQGVVDLIVPNOMLALADKSTMLAEFRYADVLLBQKGVTDLIVRGLINL 212

RESULT 4
US-09-770-509-15

;; Sequence 15, Application US/09770509
;; Publication No. US20030082657A1
;; GENERAL INFORMATION:
;; APPLICANT: Katagiri, F.
;; TITLE OF INVENTION: OOMYCETE FT5Z-MT AS A TARGET FOR

;; TITLE OF INVENTION: ANTIMICROBIAL-SPECIFIC BIOCIDES
;; FILE REFERENCE: NADII.018A
;; CURRENT APPLICATION NUMBER: US/09/770.509
;; CURRENT FILING DATE: 2001-01-25
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 15
;; LENGTH: 508
;; TYPE: PRF
;; ORGANISM: Caulobacter crescentus
US-09-770-509-15

Query Match 57.1%; Score 505; DB 9; Length 508;
Best Local Similarity 58.3%; Pred. No. 1.3e-44;
Matches 102; Conservative 28; Mismatches 45; Indels 0; Gaps 0;

QY 4 LEGEFTVAMTDCALGRSLAPHKITITGDKITGAGSKPELGKRSABQKVDIORM 63
DB 38 LEGEFTVAMTDCALGRSLAPHKITITGDKITGAGSKPELGKRSABQKVDIORM 97
QY 64 DSNMLFTTGGMGCTGGAAPVAVASVARELGILTVGVSTPFRSEGNPRTRLANAGV 123
DB 98 GAHVFTTAGMGCTGGAAPVAVASVARELGILTVGVSTPFRSEGNPRTRLANAGV 157
QY 124 AKYVDLIVPNOMLALADKSTMLAEFRYADVLLBQKGVTDLIVRGLINL 178
DB 158 QRYVDLIVPNOMLALADKSTMLAEFRYADVLLBQKGVTDLIVRGLINL 212

RESULT 5
US-09-770-509-13

;; Sequence 13, Application US/09770509
;; Publication No. US20030082657A1
;; GENERAL INFORMATION:
;; APPLICANT: Katagiri, F.
;; TITLE OF INVENTION: OOMYCETE FT5Z-MT AS A TARGET FOR
;; FILE REFERENCE: NADII.018A
;; CURRENT APPLICATION NUMBER: US/09/770.509
;; CURRENT FILING DATE: 2001-01-25
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 13
;; LENGTH: 381
;; TYPE: PRF
;; ORGANISM: Bartonella clarridgeiae
US-09-770-509-13

Query Match 56.0%; Score 496; DB 9; Length 381;
Best Local Similarity 57.7%; Pred. No. 1.4e-43;
Matches 101; Conservative 29; Mismatches 45; Indels 0; Gaps 0;

QY 4 LEGEFTVAMTDCALGRSLAPHKITITGDKITGAGSKPELGKRSABQKVDIORM 63
DB 38 LOGVFTVAMTDCALGRSLAPHKITITGDKITGAGSKPELGKRSABQKVDIORM 97
QY 64 DSNMLFTTGGMGCTGGAAPVAVASVARELGILTVGVSTPFRSEGNPRTRLANAGV 123
DB 98 DSHVFTTAGMGCTGGAAPVAVASVARELGILTVGVSTPFRSEGNPRTRLANAGV 157
QY 124 AKYVDLIVPNOMLALADKSTMLAEFRYADVLLBQKGVTDLIVRGLINL 178
DB 158 QRYVDLIVPNOMLALADKSTMLAEFRYADVLLBQKGVTDLIVRGLINL 212

RESULT 6
US-09-815-242-10639

;; Sequence 10639, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: zyskind, Judith W.

```
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: Prokaryotes
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10639
LENGTH: 410
TYPE: PRF
ORGANISM: Enterococcus faecalis
US-09-815-242-10639
```

```
Query Match      55.9%; Score 495; DB 10; Length 410;
Best Local Similarity 57.1%; Pred. No. 1,1e-43;
Matches 100; Conservative 28; Mismatches 47; Indels 0; Gaps 0;
```

```
QY 4 LGEVEFIVANTDQALGRSLAPKHTITGDKITGKLGAGSKPELGRSAEQKVDIORMQ 63
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 36 VKGVEFIVANTDQALGRSLAPKHTITGDKITGKLGAGSKPELGRSAEQKVDIORMQ 95
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 64 DSNMFLITGGMGGTGTGAAPVAVASVARELGILTVGVVSTPFSEGRPNRRLANAGVEL 123
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 96 GADMIFITAGMGGTGTGAAPVAVAKIKELGALTGVGVTRPFSEGRPNRRLANAGVEL 155
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 124 AKYVDLIYVNPONLALADKSTMLFAFRYADVLLEGVGVVDLIVRGLINL 178
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 156 KENVDTLLITSNRRLLEVVDKTPMLEAFREADVNLQGVGIGIDLTAPGYVNL 210
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 7
US-09-815-242-4909
Sequence 4909, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: Prokaryotes
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
```

```
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4909
LENGTH: 411
TYPE: PRF
ORGANISM: Enterococcus faecalis
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(411)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-4909
```

```
Query Match      55.9%; Score 495; DB 10; Length 411;
Best Local Similarity 57.1%; Pred. No. 1,1e-43;
Matches 100; Conservative 28; Mismatches 47; Indels 0; Gaps 0;
```

```
QY 4 LGEVEFIVANTDQALGRSLAPKHTITGDKITGKLGAGSKPELGRSAEQKVDIORMQ 63
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 36 VKGVEFIVANTDQALGRSLAPKHTITGDKITGKLGAGSKPELGRSAEQKVDIORMQ 95
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 64 DSNMFLITGGMGGTGTGAAPVAVASVARELGILTVGVVSTPFSEGRPNRRLANAGVEL 123
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 96 GADMIFITAGMGGTGTGAAPVAVAKIKELGALTGVGVTRPFSEGRPNRRLANAGVEL 155
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 124 AKYVDLIYVNPONLALADKSTMLFAFRYADVLLEGVGVVDLIVRGLINL 178
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 156 KENVDTLLITSNRRLLEVVDKTPMLEAFREADVNLQGVGIGIDLTAPGYVNL 210
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 8
US-09-770-509-16
Sequence 16, Application US/09770509
Patent No. US20030082657A1
GENERAL INFORMATION:
APPLICANT: Katagiri, F.
TITLE OF INVENTION: COMYCETE FTSZ-MT AS A TARGET FOR
FILE REFERENCE: NADIT 018A
CURRENT APPLICATION NUMBER: US/09/770,509
PRIOR FILING DATE: 2001-01-25
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 407
TYPE: PRF
ORGANISM: Cyanidioscyzon merolae
US-09-770-509-16
```

```
Query Match      55.4%; Score 490; DB 9; Length 407;
Best Local Similarity 53.4%; Pred. No. 3.6e-43;
Matches 95; Conservative 34; Mismatches 49; Indels 0; Gaps 0;
```

```
QY 1 ASQLEGEFIVANTDQALGRSLAPKHTITGDKITGKLGAGSKPELGRSAEQKVDIORMQ 60
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 117 ASSLPVEFIVANTDQALGRSLAPKHTITGDKITGKLGAGSKPELGRSAEQKVDIORMQ 176
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 MADSMLFITGGMGGTGTGAAPVAVASVARELGILTVGVVSTPFSEGRPNRRLANAGV 120
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 177 EFRGVALLPVTAGMGGTGTGAAPVAVAKIKELGALTGVGVTRPFSEGRPNRRLANAGV 236
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 KELAKYVDLIYVNPONLALADKSTMLFAFRYADVLLEGVGVVDLIVRGLINL 178
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 237 VELTEHVDTMLVNPONLALADKSTMLFAFRYADVLLEGVGVVDLIVRGLINL 294
   :||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 9
US-09-712-363-221

; Sequence 221, Application US/09712363
; Patent No. US20020164588A1

; GENERAL INFORMATION:

; APPLICANT: Eisenberg, David

; APPLICANT: Rotstein, Sergio H.

; APPLICANT: Marcotte, Edward M.

; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND

; FILE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS

; FILE REFERENCE: 07419-032001

; CURRENT APPLICATION NUMBER: US/09/712,363

; CURRENT FILING DATE: 2000-11-13

; PRIOR APPLICATION NUMBER: PCT/US00/02246

; PRIOR FILING DATE: 2000-01-28

; PRIOR APPLICATION NUMBER: 60/179,531

; PRIOR FILING DATE: 2000-02-01

; PRIOR APPLICATION NUMBER: 60/117,844

; PRIOR FILING DATE: 1999-01-29

; PRIOR APPLICATION NUMBER: 60/118,206

; PRIOR FILING DATE: 1999-02-01

; PRIOR APPLICATION NUMBER: 60/126,593

; PRIOR FILING DATE: 1999-03-26

; PRIOR APPLICATION NUMBER: 60/134,093

; PRIOR FILING DATE: 1999-05-14

; PRIOR APPLICATION NUMBER: 60/134,092

; PRIOR FILING DATE: 1999-05-14

; PRIOR APPLICATION NUMBER: 60/165,124

; PRIOR FILING DATE: 1999-11-12

; PRIOR APPLICATION NUMBER: 60/165,086

; PRIOR FILING DATE: 1999-11-12

; NUMBER OF SEQ ID NOS: 292

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 221

; LENGTH: 379

; TYPE: PRT

; ORGANISM: Mycobacterium tuberculosis

; US-09-712-363-221

; Query Match

; Best Local Similarity 54.8%; Score 485; DB 9; Length 379;

; Matches 95; Conservative 35; Mismatches 45; Indels 0; Gaps 0;

; Db

; 4 LEVEFIVANTDQALGRSLAPHKITLTKDITKGLGAGSKPELGKRSACQKVDIORMQ 63

; 32 LKGEFIVANTDQALGRSLAPHKITLTKDITKGLGAGSKPELGKRSACQKVDIORMQ 91

; QY

; 64 DSNMFLITGGMGGCTGCAAPVAVASARELGLTVGVSTPFRSEGNPRTLANAGV 123

; Db

; 92 GADMFVAVAGGGGCTGCAAPVAVASARELGLTVGVSTPFRSEGNPRTLANAGV 151

; QY

; 124 AKYVDLIVPNOMLALADSKSTMLAFRYADVLLGKGVTDLIVRPGILNL 178

; Db

; 152 RESCTLIIVIPNDRLQMGDAVSLMDAFSADEVLLNGVGTDLITPGLIN 206

; RESULT 10

; US-09-770-509-11

; Sequence 11, Application US/09770509

; Publication No. US20030082657A1

; GENERAL INFORMATION:

; APPLICANT: Katagiri, F.

; APPLICANT: Oomycete FT52-MT AS A TARGET FOR

; TITLE OF INVENTION: ANTIMICROBIAL-SPECIFIC BIOCIDES

; FILE REFERENCE: NADII.018A

; CURRENT APPLICATION NUMBER: US/09/770,509

; CURRENT FILING DATE: 2001-01-25

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11

; LENGTH: 583

; TYPE: PRT

; ORGANISM: Agrobacterium tumefaciens

US-09-770-509-11

Query Match 54.5%; Score 482; DB 9; Length 583;

Best Local Similarity 57.1%; Pred. No. 4,1e-42;

Matches 100; Conservative 26; Mismatches 49; Indels 0; Gaps 0;

; Db

; 4 LEVEFIVANTDQALGRSLAPHKITLTKDITKGLGAGSKPELGKRSACQKVDIORMQ 63

; Db

; 38 LQGDVFNANTDQALGRSLAPHKITLTKDITKGLGAGSKPELGKRSACQKVDIORMQ 97

; QY

; 64 DSNMFLITGGMGGCTGCAAPVAVASARELGLTVGVSTPFRSEGNPRTLANAGV 123

; Db

; 98 GTHMCFYTAGGGCTGCAAPVAVASARELGLTVGVSTPFRSEGNPRTLANAGV 157

; QY

; 124 AKYVDLIVPNOMLALADSKSTMLAFRYADVLLGKGVTDLIVRPGILNL 178

; Db

; 158 OKSVDTLIVPNOMLALADSKSTMLAFRYADVLLGKGVTDLIVRPGILNL 212

; RESULT 11

; US-09-770-509-10

; Sequence 10, Application US/09770509

; Publication No. US20030082657A1

; GENERAL INFORMATION:

; APPLICANT: Katagiri, F.

; APPLICANT: Oomycete FT52-MT AS A TARGET FOR

; TITLE OF INVENTION: ANTIMICROBIAL-SPECIFIC BIOCIDES

; FILE REFERENCE: NADII.018A

; CURRENT APPLICATION NUMBER: US/09/770,509

; CURRENT FILING DATE: 2001-01-25

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10

; LENGTH: 420

; TYPE: PRT

; ORGANISM: Phytophthora infestans

; US-09-770-509-10

; Query Match

; Best Local Similarity 54.4%; Score 481; DB 9; Length 420;

; Matches 92; Conservative 39; Mismatches 47; Indels 0; Gaps 0;

; Db

; 1 ASOLEVEFIVANTDQALGRSLAPHKITLTKDITKGLGAGSKPELGKRSACQKVDIORMQ 60

; Db

; 129 ARGQGEFIVANTDQALGRSLAPHKITLTKDITKGLGAGSKPELGKRSACQKVDIORMQ 188

; QY

; 61 MLDSONMFLITGGMGGCTGCAAPVAVASARELGLTVGVSTPFRSEGNPRTLANAGV 120

; Db

; 189 RVQGANMFYTAGGGCTGCAAPVAVASARELGLTVGVSTPFRSEGNPRTLANAGV 248

; QY

; 121 KELAKYVDLIVPNOMLALADSKSTMLAFRYADVLLGKGVTDLIVRPGILNL 178

; Db

; 249 AELKDSVDLIVPNOMLALADSKSTMLAFRYADVLLGKGVTDLIVRPGILNL 306

; RESULT 12

; US-09-815-242-5770

; Sequence 5770, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: Elittra 011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21


```
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 5770
;; LENGTH: 390
;; TYPE: PRT
;; ORGANISM: Staphylococcus aureus
US-09-815-242-5770
```

```
Query Match          54.2%; Score 480; DB 10; Length 390;
Best Local Similarity 54.3%; Pred. No. 3.8e-42;
Matches 95; Conservative 33; Mismatches 47; Indels 0; Gaps 0;
```

```
OY 4 LEVEFIYANTDQALGRSLAPHKITLTKGDKGAGSKPELGRSAEQKVDIQRLQ 63
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35 MNVEFIINTDQALNLSKASKIQIGKILRGLGAGANPELGRKAAESRQIEDAIQ 94
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 64 DSNMLFTGGMGGCTGCGAAPVAVASVARELIGITVGVSTPRSEGNPRTLANAGVKEL 123
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 95 GADVFYVSGMGCGTGTGAAPVAVAKIEMGALTGVVTRPFSEGRKRQTOAAAGVEAM 154
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 124 AKYVDTLIVPNQNLALADKSTMLAEFRYADVLEGVKGVTDLIYRGLINL 178
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 KAAVDTLIVPNDRLLDIYDKSTPMKEAFKADNVLKGVGIGISDLIAVSGEYVL 209
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 13
US-09-815-242-12723
;; Sequence 12723, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Travick, John D.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE REFERENCE: EXTRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
```

```
;; SEQ ID NO 12723
;; LENGTH: 392
;; TYPE: PRT
;; ORGANISM: Staphylococcus aureus
US-09-815-242-12723
```

```
Query Match          54.2%; Score 480; DB 10; Length 392;
Best Local Similarity 54.3%; Pred. No. 3.8e-42;
Matches 95; Conservative 33; Mismatches 47; Indels 0; Gaps 0;
```

```
OY 4 LEVEFIYANTDQALGRSLAPHKITLTKGDKGAGSKPELGRSAEQKVDIQRLQ 63
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35 MNVEFIINTDQALNLSKASKIQIGKILRGLGAGANPELGRKAAESRQIEDAIQ 94
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 64 DSNMLFTGGMGGCTGCGAAPVAVASVARELIGITVGVSTPRSEGNPRTLANAGVKEL 123
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 95 GADVFYVSGMGCGTGTGAAPVAVAKIEMGALTGVVTRPFSEGRKRQTOAAAGVEAM 154
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 124 AKYVDTLIVPNQNLALADKSTMLAEFRYADVLEGVKGVTDLIYRGLINL 178
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 KAAVDTLIVPNDRLLDIYDKSTPMKEAFKADNVLKGVGIGISDLIAVSGEYVL 209
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 14
US-09-770-509-12
;; Sequence 12, Application US/09770509
;; Publication No. US20030082657A1
;; GENERAL INFORMATION:
;; APPLICANT: Katagiri, F.
;; TITLE OF INVENTION: COMYCETE FTS2-MT AS A TARGET FOR
;; FILE REFERENCE: NADII.018A
;; CURRENT APPLICATION NUMBER: US/09/770,509
;; CURRENT FILING DATE: 2001-01-25
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 12
;; LENGTH: 590
;; TYPE: PRT
;; ORGANISM: Sinorhizobium meliloti
US-09-770-509-12
```

```
Query Match          53.7%; Score 475; DB 9; Length 590;
Best Local Similarity 55.4%; Pred. No. 2.2e-41;
Matches 97; Conservative 29; Mismatches 49; Indels 0; Gaps 0;
```

```
OY 4 LEVEFIYANTDQALGRSLAPHKITLTKGDKGAGSKPELGRSAEQKVDIQRLQ 63
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38 LOGVDFVANTDQALNLSKASKIQIGKILRGLGAGANPELGRKAAESRQIEDAIQ 97
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 64 DSNMLFTGGMGGCTGCGAAPVAVASVARELIGITVGVSTPRSEGNPRTLANAGVKEL 123
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98 GTHMCFYTAGMGCGTGTGAAPVAVAKIEMGALTGVVTRPFSEGRKRQTOAAAGVEAM 157
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 124 AKYVDTLIVPNQNLALADKSTMLAEFRYADVLEGVKGVTDLIYRGLINL 178
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 OKSVDTLIVPNQNLALADKSTMLAEFRYADVLEGVKGVTDLIYRGLINL 212
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 15
US-09-754-608-2
;; Sequence 2, Application US/09754608
;; Patent No. US20020004580A1
;; GENERAL INFORMATION:
;; APPLICANT: Fueyo, Joanna Lynn
;; APPLICANT: Lonetto, Michael A.
;; APPLICANT: Pearce, Kenneth
;; TITLE OF INVENTION: fts2
;; FILE REFERENCE: GM10068
;; CURRENT APPLICATION NUMBER: US/09/754,608
;; CURRENT FILING DATE: 2001-01-04
;; PRIOR APPLICATION NUMBER: 09/120,426
;; PRIOR FILING DATE: 1998-07-22
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2003, 08:33:30 ; Search time 271 Seconds

(without alignments)
423.478 Million cell updates/sec

Title: US-09-770-509-2

Perfect score: 885
Sequence: 1 ASOLEGVEFIVANTDCQALG.....LLEGVKTVDLVRGLINL 178

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Pending_Patents_AA_Main: *
1: /cgn2_6/ptodata/1/paa/PCYUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	885	100.0	178	21	US-09-770-509-2
2	527	59.5	336	21	US-09-791-537-32478
3	527	59.5	401	21	US-09-770-509-17
4	527	59.5	401	21	US-09-791-537-130384
5	520	58.8	370	21	US-09-791-537-1723
6	518	58.5	452	21	US-09-770-509-14

7	518	58.5	452	21	US-09-791-537-69060	Sequence 69060, A
8	517	58.4	289	21	US-09-791-537-40309	Sequence 40309, A
9	515	58.2	289	21	US-09-791-537-40338	Sequence 40338, A
10	514	58.1	289	21	US-09-791-537-40306	Sequence 40306, A
11	514	58.1	289	21	US-09-791-537-40308	Sequence 40308, A
12	514	58.1	289	21	US-09-791-537-40312	Sequence 40312, A
13	510	57.6	398	21	US-09-791-537-29481	Sequence 29481, A
14	506	57.2	399	21	US-09-791-537-105171	Sequence 105171, A
15	506	57.2	399	21	US-09-791-537-132641	Sequence 132641, A
16	506	57.2	404	21	US-09-791-537-65811	Sequence 65811, A
17	506	57.2	404	21	US-09-791-537-132535	Sequence 132535, A
18	505	57.1	508	21	US-09-770-509-15	Sequence 15, Appl
19	505	57.1	508	21	US-09-791-537-74160	Sequence 74160, A
20	499	56.4	590	21	US-09-791-537-22908	Sequence 22908, A
21	498	56.3	382	21	US-09-791-537-10179	Sequence 10179, A
22	496	56.0	381	21	US-09-770-509-13	Sequence 13, Appl
23	496	56.0	410	21	US-09-791-537-125714	Sequence 125714, A
24	495	55.9	581	21	PCT-US02-03987-10639	Sequence 10639, A
25	495	55.9	410	22	US-09-815-242-10639	Sequence 10639, A
26	495	55.9	410	24	US-10-072-851-10639	Sequence 10639, A
27	495	55.9	411	1	PCT-US02-03987-4909	Sequence 4909, Ap
28	495	55.9	411	22	US-09-815-242-4909	Sequence 4909, Ap
29	495	55.9	411	24	US-10-072-851-4909	Sequence 4909, Ap
30	495	55.9	411	27	US-09-791-537-4678	Sequence 4678, Ap
31	495	55.9	435	15	US-09-134-000-4658	Sequence 4658, Ap
32	494	55.8	414	15	US-09-107-532-7295	Sequence 7295, Ap
33	494	55.8	414	15	US-09-107-532-7295	Sequence 7295, Ap
34	494	55.8	592	21	US-09-791-537-132634	Sequence 132634, A
35	493	55.7	372	21	US-09-791-537-1738	Sequence 1738, Ap
36	492	55.6	413	21	US-09-791-537-34221	Sequence 34221, A
37	491	55.5	351	21	US-09-791-537-11835	Sequence 11835, A
38	490	55.4	407	21	US-09-770-509-16	Sequence 16, Appl
39	489	55.3	407	21	US-09-791-537-114519	Sequence 114519, A
40	489	55.3	412	21	US-09-791-537-77265	Sequence 77265, A
41	488	55.1	419	21	US-09-791-537-82315	Sequence 82315, A
42	486.5	55.0	318	21	US-09-791-537-31564	Sequence 31564, A
43	486	54.9	399	21	US-09-791-537-1722	Sequence 1722, Ap
44	486	54.9	399	21	US-09-791-537-29477	Sequence 29477, A
45	486	54.9	402	21	US-09-791-537-132563	Sequence 132563, A

ALIGNMENTS

RESULT 1
US-09-770-509-2
; Sequence 2, Application US/09770509
; GENERAL INFORMATION:
; APPLICANT: Katagiri, F.
; TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
; TITLE OF INVENTION: ANTIMICROBIAL-SPECIFIC BIODES
; FILE REFERENCE: NADII-018A
; CURRENT APPLICATION NUMBER: US/09/770,509
; CURRENT FILING DATE: 2001-01-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Phytophthora infestans
US-09-770-509-2

QY	1	ASOLEGVEFIVANTDCQALGRLAPKHTITGKITKGLGSKPELGRSAEOQVIO	60
DB	1	ASOLEGVEFIVANTDCQALGRLAPKHTITGKITKGLGSKPELGRSAEOQVIO	60
QY	61	MLQDSNMLFITGGMGGCTCTGAAPVAVASVARELGIITLVGVSTPFRSEGNRRRLANAGV	120
DB	61	MLQDSNMLFITGGMGGCTCTGAAPVAVASVARELGIITLVGVSTPFRSEGNRRRLANAGV	120

OY 121 KELATVDTLIVPQNQLALADKSTMTLEAFRYADVLLGEGKVTDLIVRPGILNL 178
 |||||
D6 121 KELATVDTLIVPQNQLALADKSTMTLEAFRYADVLLGEGKVTDLIVRPGILNL 178

RESULT 2

```

US-09-791-537-32478
: Sequence 32478, Application US/09791537
:
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Dargatz, Joseph
:
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
:
: TITLE OF INVENTION: METHODS OF USE THEREOF
:
: FILE REFERENCE: 261/210
:
: CURRENT APPLICATION NUMBER: US/09/791,537
:
: CURRENT FILING DATE: 2001-02-22
:
: NUMBER OF SEQ ID NOS: 153055
:
: SOFTWARE: PatentIn version 3.0
:
: SEQ ID NO 32478
:
: LENGTH: 336
:
: TYPE: prt
:
: ORGANISM: Zymomonas mobilis
:
: US-09-791-537-32478

```

Query Match	59.58	Score 527	DB 21	Length 336
Best Local Similarity	59.68	Pred. No. 4.4e-50		
Matches 106	Conservative 29	Mismatches 43	Indels 0	Gaps 0

```

QY 1 ASOLEGVEFIYANTDQOALGRSLAHPKHTLKGDIKGLDAGSKPELGRKBAEOQVDIOR 60
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 38 ASGQGVDFIYANTDQALINSPAEORIOLGPTTIOGLASNPBEVKGAAEETIEOIE 97
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 61 MLODSNMLEITGGMGSGTGAAPVYASVARELGILTYGVSNRPFSEGNPTRIANAGV 12
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 98 ALEGRRMCFIANGMGSGTGTGAAPYIAKARDGILTYGVYIKPFMEFEGRRRRASASGI 157
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 121 KEKARYDVTLLIYVPMONLALADKSTTLMLEAFERADVLELEGKGYTDLIVRGLINL 178
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 158 EELQKHVDTLIYVPMONLEFIANPNTTFEYQOAMDEVLDQGVKGTITDLIVCGGLINL 215
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```

RESULT 3

```

US-09-770-509-17
: Sequence 17, Application US/09770509
: GENERAL INFORMATION:
:   APPLICANT: Katagiri, F.
:   TITLE OF INVENTION: OOMYCETE FTSG-MT AS A TARGET FOR
:   TITLE OF INVENTION: ANTIMICROBIAL-SPECIFIC BIOCIDES
:   FILE REFERENCE: NADIT. 018A
:   CURRENT APPLICATION NUMBER: US/09/770,509
:   CURRENT FILING DATE: 2001-01-25
:   NUMBER OF SEQ ID NOS: 32
:   SOFTWARE: FastsEq for Windows Version 4.0
:   SEQ ID NO 17
:   LENGTH: 401
:   TYPE: PR1
:   ORGANISM: Mallomonas splendens
US-09-770-509-17

```

Query Match	59.5%	Score 527	DB 21	Length 401
Best Local Similarity	58.4%	Pred. No. 5.8e-50		
Matches 104	Conservative 27	Mismatches 47	Indels 0	Gaps 0

QY	1	ASOLEGVEIVANTCCOALGRSLAPKRLTKGRITGLAGSGAPPELGRKREAOQVJOR	60
Db	100	ARKLSGVEFVCANTDAQHLSTCLTEKRLDGLKRESTGLGGANPEESGRRAAESKEIAR	150
QY	61	MLQDSNNMLFITGGKGGCTGTGAAPVYASARELGITITGVYSSPFSSEGNRRIRLAAQV	120
Db	160	YIADANMFITRAGKGGGTGTGAAPVYAEQMEKDIITLVAAYVTKPFSEGHRRALAEGL	210

```
Qy      121 KELAKYDVTLIIVPÑÖNLALADKSTMTLEAFRYADVLTLEGKGVTDLIVRPGLINL 178
        : | |||||::||||: | : ||: || ||||: || ||| : ||||: |||||
Db      220 RSELDVRDTLLIIPÑÖNIFKLINASTSMADAFGLADDILLAGVKSITDLIMVRPGLINL 277
```

RESULT 4

```

US-09-791-537-130384
: Sequence 130384, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debez, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 130384
: LENGTH: 401
: TYPE: prt
: ORGANISM: Mallomonas splendens
: US-09-791-537-130384

```

Query Match	59.5%	Score	527	DB	21	length	401
Best Local Similarity	58.4%	Pred. No.	5.8e-50				
Matches	104	Conservative	27	Mismatches	47	Indels	0
						Gaps	0

OY 1 ASOLEGVEIYANTCOALGRSLAPKRTITLGRITGLAGSKPELGAKREAEOKVOR 60
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Db 100 ARKLSTGEVFCANTDAHLSCTLETKRLQJGKSTQGLGGCAGPESGRRAAESKEETAR 158
 OY 61 MLODSNMFLITGMGGGTCTGAAPVYASVARBELGITLVGVSNPFPSSEGNFRIARAGV 120
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Db 160 YIADANNPFITGMGGGTCTGAAPVYAEBCEMEDILTVANVTNPFSEEGKHARLANEGI 219
 OY 121 KELAYVDLLIVPNOMNLADKSTTMEAFRYADDVLLEGYKGATDLYVRGLNL 178
 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 Db 220 RSLERAVPDLLLLIPONIKPLINASTSMADPAELADIDILLAGKSTDLDWRRGLNL 277

RESULT 5
US-09-75

```

: Sequence 1723, Application US/09791537
:
: GENERAL INFORMATION:
:
: APPLICANT: Bionomix, Inc.
:
: APPLICANT: Debe, Derek
:
: APPLICANT: Danzer, Joseph
:
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
:
: TITLE OF INVENTION: METHODS OF USE THEREOF
:
: FILE REFERENCE: 261/210
:
: CURRENT APPLICATION NUMBER: US/09/791,537
:
: CURRENT FILING DATE: 2001-02-22
:
: NUMBER OF SEQ ID NOS: 153055
:
: SOFTWARE: PatentIn version 3.0
:
: SEQ ID NO 1723
:
: LENGTH: 370
:
: TYPE: prt
:
: ORGANISM: Clostridium lentocellum
:
: US-09-791-537-1723

```

Query Match	58.8%	Score 520	DB 21	Length 370
Best Local Similarity	58.9%	Pred.No. 3.2e-49		
Matches 103; Conservative	30	Mismatches 42	Indels 0	Gaps 0

```

0Y 4 LEGVEFIATNDCOLLRSLAPHKITLGDITDITGLAGSPPELGKRSASEOQYIDQML 63
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35 LEGVEFIYVNTDHOALRRSGAPAKIOIGKMTDGLGAGAPPELGTSTSAESRELLTAIK 94
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

0Y 64 DSNMFLITGGMGCGTCTGAAPVYVASARELGIITGVYSPFESGCPNTRLNAGYKEL 12
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 95 GADMLFIITAGMGCGGTGAAPYIASIASTAKEGIIITGVYVTKPSEFEGKKRIRINERKGAEL 15
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Oy      124 AKYVDLIVPNQNLALADKSTYMEAFRYADVDLLEGVKGYTDLIVRGLINTL 178
          : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      155 KQNVDFLVIIPNDKILQVIDKKTMYDAFSKADVDLQGVGVGTDLISNPGIINTL 209

```

```

RESULT 6
US-09-770-509-14
: Sequence 14, Application US/09770509
: GENERAL INFORMATION:
: APPLICANT: Katagiri, F.
: TITLE OF INVENTION: COMCYETE FT52-MT AS A TARGET FOR
: TITLE OF INVENTION: ANTIMICROBIAL-SPECIFIC BIODES
: FILE REFERENCE: NADII 018A
: CURRENT APPLICATION NUMBER: US/09/770,509
: CURRENT FILING DATE: 2001-01-25
: NUMBER OF SEQ ID NOS: 32
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 14
: LENGTH: 452
: TYPE: PRF
: ORGANISM: Rickettsia prowazekii
: US-09-770-509-14

```

	Query Match	58.5%	Score 518;	DB 21;	Length 452;
	Best Local Similarity	55.9%	Pred. No. 7.2e-49;		
	Matches	99;	Conservative	38;	Mismatches 40; Indels 0; Gaps 0
QY	2	SQLEGEVETVANTDQALGRLSLAPKHTLTKDKDTKGLAGSGSPELCKRSAGEOQKYDIORM	61		
		: : - : - : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :			
Db	36	ANLOGANFVVANTDAQSLEHSLCIKIDLGAVSTRTGLGASGAPRGALGAJSENNIRSS	95		
QY	62	LQDSNMLEFITGGMGCGTCTGCAPPVVASVARELGITLVGVSTPFPSSEGNKRRLNAGAK	121		
		: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :			
Db	96	LENENMVFITAGMGCGTGTCGAPPIIARLAKEIGITLVGVTPRFHEGEGHRMKATADGILI	155		
QY	122	ELAKYVDPLIVPQNLLADSKSTTMLAEFRAYADDVLLEGKGYTDILVRGTLNL	178		
		: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :			
Db	156	ELQGFVDILVIYPNQNLRIANEQTTFADAFRMADVDLAGRGVATDLIMPGGLNL	212		

```

RESULT 7
US-09-791-537-69060
; Sequence 69060, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 69060
; LENGTH: 452
; TYPE: PRN
; ORGANISM: Rickettsia prowazekii
; US-09-791-537-69060

```

[illegible]

```
QY      122 ELAKYVDLIVPNQNLALADKSTMLAEFRADVLLLEGVKGVMDLIVRPELINL 178
      ||:::|||||:::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      156 ELQGFVDLIVPNQNLFRIANEQTFADAFKMAVDVLHAGVRGVYDLMIMPELINL 212
```

```

RESULT 8
US-09-791-537-40309
: Sequence 40309, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS THEREOF
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 40309
: LENGTH: 289
: TYPE: PRF
: ORGANISM: Wolbachia sp
: US-09-791-537-40309

```

	Query Match	58.4%	Score 517;	DB 21;	Length 289;
	Best Local Similarity	56.6%	Pred. NO. 4.8e-49;		
	Matches 107;	Conservative 29;	Mismatches 41;	Indels 12;	Gaps 1.
QY	2	SQLEGEVFIYANTDCAQLGRSLAHPKITYLGDITKGLAGAGSKPELGKRSAEQKYDIORM	61		
Db	1	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	61		
	2	SNLDGVFVVYVANTDAOLAEKSLCDKKIQDLGINTRLGLGAGALPDVGKAESIDEIMEH	61		
QY	62	LQDSNMIFETMGSGGGCTGTGAAPVAVASAREL-----GILTVGYVSITPFRSEEG	109		
Db	62	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	121		
	62	IKDSHMFLTRGMGGTGTGAAPVIAKARARAAYVDRAEKKITLVGWATREFGEGE	121		
QY	110	PNRTLRNAGVKEELAKYVDTLIVVPNONLLADLSTMLEAFRYADVDDLEGVGYDDL	169		
Db	122	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	181		
	122	VRRRIARIAELEHELOTKYVDTLIVIPNQNLRFIANEKTTFSDAFKLADVNLHGIRGYIDL	181		
QY	170	IYRGLINL 178			
Db	182	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	190		
	182	MYMFGLINL 190			

```

RESULT 9
US-09-791-537-40338
: Sequence 40338, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Biomomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 40338
: LENGTH: 289
: TYPE: PRT
: ORGANISM: Wolbachia sp
: US-09-791-537-40338

```

```

Query Match      58.2%  Score 515; DB 21, Length 289;
Best Local Similarity 56.6%  Pred. NO, 8e-49;
Matches 107; Conservative 28; Mismatches 42; Indels 12; Gaps 1

```

QY 62 LODSNMFLITGGMGGCTGCAAPVAVASVAREL-----GILTVGVSTPFRSE 109
DB 62 IRDSHMLFTTAGMGCGTGTGAAPVIAKAAAREARAVVKKGAKKEKILTVGVYTKPFGE 121
QY 110 PNRTRLANAGVKELEAKYVDLTVVNPONMLALADKSTMLAEFRYADVLLBEGKVDTL 169
DB 122 VRRMRTAELGLEELQKYVDLTVIPNOMLFRIANEKTTFADAFQDLADNVHLHIGIRGVTDL 181
QY 170 IVRPGLINTL 178
DB 182 MIMPGILINTL 190

RESULT 10
US-09-791-537-40306
; Sequence 40306, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40306
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Wolbachia sp
US-09-791-537-40306

Query Match
Best Local Similarity 58.1%; Score 514; DB 21; Length 289;
Matches 106; Conservative 29; Mismatches 42; Indels 12; Gaps 1;

QY 2 SOLEGEVEFIYANTDCCALGRSLAPHKITLGDITKGLGAGSKPELGKRSAEQKVDIORM 61
DB 2 SNLOGVNFVYANTDQALEKSLCDKKTQIGINTLTKGAGALPDIGKAAEESIDEIMEH 61
QY 62 LODSNMFLITGGMGGCTGCAAPVAVASVAREL-----GILTVGVSTPFRSE 109
DB 62 IRDSHMLFTTAGMGCGTGTGAAPVIAKAAAREARAVVKKGAKKEKILTVGVYTKPFGE 121
QY 110 PNRTRLANAGVKELEAKYVDLTVVNPONMLALADKSTMLAEFRYADVLLBEGKVDTL 169
DB 122 VRRMRTAELGLEELQKYVDLTVIPNOMLFRIANEKTTFADAFQDLADNVHLHIGIRGVTDL 181
QY 170 IVRPGLINTL 178
DB 182 MIMPGILINTL 190

RESULT 11
US-09-791-537-40308
; Sequence 40308, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40308
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Wolbachia sp

US-09-791-537-40308

Query Match
Best Local Similarity 58.1%; Score 514; DB 21; Length 289;
Matches 106; Conservative 29; Mismatches 42; Indels 12; Gaps 1;

QY 2 SOLEGEVEFIYANTDCCALGRSLAPHKITLGDITKGLGAGSKPELGKRSAEQKVDIORM 61
DB 2 SNLOGVNFVYANTDQALEKSLCDKKTQIGINTLTKGAGALPDIGKAAEESIDEIMEH 61
QY 62 LODSNMFLITGGMGGCTGCAAPVAVASVAREL-----GILTVGVSTPFRSE 109
DB 62 IRDSHMLFTTAGMGCGTGTGAAPVIAKAAAREARAVVKKGAKKEKILTVGVYTKPFGE 121
QY 110 PNRTRLANAGVKELEAKYVDLTVVNPONMLALADKSTMLAEFRYADVLLBEGKVDTL 169
DB 122 VRRMRTAELGLEELQKYVDLTVIPNOMLFRIANEKTTFADAFQDLADNVHLHIGIRGVTDL 181
QY 170 IVRPGLINTL 178
DB 182 MIMPGILINTL 190

RESULT 12
US-09-791-537-40312
; Sequence 40312, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40312
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Wolbachia sp
US-09-791-537-40312

Query Match
Best Local Similarity 58.1%; Score 514; DB 21; Length 289;
Matches 106; Conservative 29; Mismatches 42; Indels 12; Gaps 1;

QY 2 SOLEGEVEFIYANTDCCALGRSLAPHKITLGDITKGLGAGSKPELGKRSAEQKVDIORM 61
DB 2 SNLOGVNFVYANTDQALEKSLCDKKTQIGINTLTKGAGALPDIGKAAEESIDEIMEH 61
QY 62 LODSNMFLITGGMGGCTGCAAPVAVASVAREL-----GILTVGVSTPFRSE 109
DB 62 IRDSHMLFTTAGMGCGTGTGAAPVIAKAAAREARAVVKKGAKKEKILTVGVYTKPFGE 121
QY 110 PNRTRLANAGVKELEAKYVDLTVVNPONMLALADKSTMLAEFRYADVLLBEGKVDTL 169
DB 122 VRRMRTAELGLEELQKYVDLTVIPNOMLFRIANEKTTFADAFQDLADNVHLHIGIRGVTDL 181
QY 170 IVRPGLINTL 178
DB 182 MIMPGILINTL 190

RESULT 13
US-09-791-537-29481
; Sequence 29481, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS

FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791.537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 29481
LENGTH: 398
TYPE: PR1
ORGANISM: Wolbachia sp
US-09-791-537-29481

Query Match 57.6%; Score 510; DB 21; Length 398;
Best Local Similarity 56.1%; Pred. No. 4.7e-48;
Matches 106; Conservative 29; Mismatches 42; Indels 12; Gaps 1;

QY 2 SOLEGEVFIYANDCOALGRSLAPHKITLTKGDKGAGSKPELGKRSADQKVDIORM 61
DB 36 SNLQGVNFVANVDQALERSLCKDKIQLGINTLTKGAGALPDVGKAAEESIDEIMH 95
QY 62 LODSNMLFTGGMGGCTGGAAPVYASVAREL-----GILTVGVSTPFRSEG 109
DB 96 IKDSHMLFTAGMGCTGGAAPVYAKARARAANDRAKKEKILTVGVTKRFGREG 155
QY 110 PNRRLANAGVKELAKYVDLLIVPNONLALADKSTTMLFAFRYADDVLLGKGVYDLE 169
DB 156 VRHMPRIELGLELQKYVDLLIVPNONLFRANKEKTFSDAFKLADVHLHIGRYVDL 215
QY 170 IYVRGLINL 178
DB 216 VMMPGLINL 224

RESULT 14

US-09-791-537-105171
Sequence 105171, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Blononix, Inc.
APPLICANT: Danzer, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791.537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 105171
LENGTH: 399
TYPE: PR1
ORGANISM: Borrelia burgdorferi
US-09-791-537-105171

Query Match 57.2%; Score 506; DB 21; Length 399;
Best Local Similarity 58.7%; Pred. No. 1.3e-47;
Matches 101; Conservative 27; Mismatches 44; Indels 0; Gaps 0;

QY 7 VERIVANTDQALGRSLAPHKITLTKGDKGAGSKPELGKRSADQKVDIORMLODSN 66
DB 47 VERIVANTDQALQTSIAPIKIALGAKVYTAGAGKPEIGQAAEEDIDIVIRNHLGAD 106
QY 67 MLEFITGGMGGCTGGAAPVYASVARELGLTVGVSTPFRSEGNRRLANAGVKELAKY 126
DB 107 MVEFITGGMGGCTGGAAPVYAKVAKELGLTVGVTKRPFEGPKRLAEOGINLNRKS 166
QY 127 VDTLIVPNONLALADKSTTMLFAFRYADDVLLGKGVYDLLIVRGLINL 178
DB 167 VDTLIIIPNOKLTVVDKRTTIKDAFRADVDLVMGVGIAGLIIIEHGEVNI 218

RESULT 15

US-09-791-537-132641
Sequence 132641, Application US/09791537
GENERAL INFORMATION:

APPLICANT: Blononix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791.537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 132641
LENGTH: 399
TYPE: PR1
ORGANISM: Borrelia burgdorferi
US-09-791-537-132641

Query Match 57.2%; Score 506; DB 21; Length 399;
Best Local Similarity 58.7%; Pred. No. 1.3e-47;
Matches 101; Conservative 27; Mismatches 44; Indels 0; Gaps 0;

QY 7 VERIVANTDQALGRSLAPHKITLTKGDKGAGSKPELGKRSADQKVDIORMLODSN 66
DB 47 VERIVANTDQALQTSIAPIKIALGAKVYTAGAGKPEIGQAAEEDIDIVIRNHLGAD 106
QY 67 MLEFITGGMGGCTGGAAPVYASVARELGLTVGVSTPFRSEGNRRLANAGVKELAKY 126
DB 107 MVEFITGGMGGCTGGAAPVYAKVAKELGLTVGVTKRPFEGPKRLAEOGINLNRKS 166
QY 127 VDTLIVPNONLALADKSTTMLFAFRYADDVLLGKGVYDLLIVRGLINL 178
DB 167 VDTLIIIPNOKLTVVDKRTTIKDAFRADVDLVMGVGIAGLIIIEHGEVNI 218

Search completed: June 2, 2003, 08:47:09
Job time : 276 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2003, 08:41:25 ; Search time 37 Seconds
(without alignments)
1012.472 Million cell updates/sec

Title: US-09-770-509-2

Perfect score: 885

Sequence: 1 ASOLEGVERIVANTDCQALG.....LLEGVKGVTDLVPRGLNL 178

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1005382 seqs, 210457600 residues

Total number of hits satisfying chosen parameters: 1005382

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2.6/pdodata1/paa/PCIT_NEW_COMB.pep:*
2: /cgn2.6/pdodata1/paa/US06_NEW_COMB.pep:*
3: /cgn2.6/pdodata1/paa/US07_NEW_COMB.pep:*
4: /cgn2.6/pdodata1/paa/US08_NEW_COMB.pep:*
5: /cgn2.6/pdodata1/paa/US09_NEW_COMB.pep:*
6: /cgn2.6/pdodata1/paa/US10_NEW_COMB.pep:*
7: /cgn2.6/pdodata1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	510	57.6	391	6	US-10-282-122A-60624
2	507	57.3	373	6	US-10-282-122A-51618
3	506	57.2	404	6	US-10-282-122A-47081
4	505	57.1	386	6	US-10-282-122A-45663
5	505	57.1	386	6	US-10-282-122A-46846
6	505	57.1	394	6	US-10-282-122A-53331
7	495	55.3	410	6	US-10-282-122A-42469
8	495	55.9	435	5	US-09-134-000C-4658
9	495	55.9	435	5	US-09-134-000C-4658
10	494	55.8	413	6	US-10-282-122A-57592
11	494	55.8	414	6	US-10-417-884-7295
12	492	55.6	387	6	US-10-282-122A-61968
13	490	55.4	359	6	US-10-282-122A-52722
14	490	55.4	410	1	PCT-US02-36122A-84
15	490	55.4	410	1	PCT-US02-36122A-84
16	486	54.9	396	6	US-10-282-122A-54015
17	485	54.9	411	6	US-10-282-122A-62334
18	485	54.8	379	6	US-10-282-122A-64648
19	485	54.8	379	6	US-10-282-122A-64648
20	482	54.4	393	6	US-10-282-122A-71506
21	481	54.4	379	6	US-10-282-122A-63791
22	480	54.2	390	5	US-09-950-084-5515
23	480	54.2	390	6	US-10-282-122A-44008
24	480	54.2	394	6	US-10-092-411A-4834
25	480	54.2	394	6	US-10-282-122A-70732
26	472	53.3	419	6	US-10-282-122A-74113

ALIGNMENTS

27	471	53.2	398	6	US-10-282-122A-67907	Sequence 67907, A
28	468	52.9	395	6	US-10-282-122A-69702	Sequence 69702, A
29	466	52.7	421	1	PCT-US02-33727-21663	Sequence 21663, A
30	466	52.7	421	5	US-09-978-823-21663	Sequence 21663, A
31	466	52.7	421	6	US-10-057-98-21663	Sequence 21663, A
32	465.5	52.6	356	6	US-10-282-122A-63205	Sequence 63205, A
33	465	52.5	418	6	US-10-282-122A-76497	Sequence 76497, A
34	464	52.4	394	6	US-10-282-122A-66647	Sequence 66647, A
35	464	52.4	400	6	US-10-366-683-24128	Sequence 24128, A
36	464	52.4	400	6	US-10-419-128-24128	Sequence 24128, A
37	464	52.4	439	6	US-10-282-122A-74650	Sequence 74650, A
38	460	52.0	383	6	US-10-282-122A-77931	Sequence 77931, A
39	459	51.9	425	5	US-09-500-495A-6	Sequence 6, App1
40	459	51.9	434	6	US-10-282-122A-72045	Sequence 72045, A
41	455.5	51.5	391	6	US-10-282-122A-44660	Sequence 44660, A
42	453	51.2	383	6	US-10-282-122A-75194	Sequence 75194, A
43	453	51.2	383	6	US-10-282-122A-76027	Sequence 76027, A
44	453	51.2	414	6	US-10-424-599-177675	Sequence 177675, A
45	452	51.1	383	6	US-10-282-122A-42973	Sequence 42973, A

RESULT 1

US-10-282-122A-60624
Sequence 60624, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
XU, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 60624
LENGTH: 391
TYPE: PRT
ORGANISM: Listeria monocytogenes
US-10-282-122A-60624

Query Match 57.6%; Score 510; DB 6; Length 391;
Best Local Similarity 56.0%; Pred. No. 2.2e-39;
Matches 98; Conservative 35; Mismatches 42; Indels 0; Gaps 0;

QY 4 LCGVEIVANTDQALGRSLAPHKITLTKGLGAGSKPELGRSAEQOKVDIORMQ 63
DB 35 VCGVEIVANTDQALGRSLAPHKITLTKGLGAGSKPELGRSAEQOKVDIORMQ 94

QY 64 DSNMLFTIGMGGGTCTGAAPVAVSARELGILTVGVSTPERSEGNPRTLANGVKEL 123
DB 95 GSDMVEVTAGMGGGTCTGAAPVAVSARELGILTVGVSTPERSEGNPRTLANGVKEL 154

QY 124 AKYVDLIVPNONLALADKSTMLAEFRADVDLLEGVGVYDLIVRPLINL 178
DB 155 KEAVDPLIVIPNDRLQIVDKNTPELEAFREADVNLROGVGIGISDLIVPGLINL 209

RESULT 2
US-10-282-122A-51618
Sequence 51618, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 51618
LENGTH: 373
TYPE: PRT
ORGANISM: Clostridium acetobutylicum
US-10-282-122A-51618

Query Match 57.3%; Score 507; DB 6; Length 373;
Best Local Similarity 56.0%; Pred. No. 3.9e-39;
Matches 98; Conservative 35; Mismatches 42; Indels 0; Gaps 0;

QY 4 LCGVEIVANTDQALGRSLAPHKITLTKGLGAGSKPELGRSAEQOKVDIORMQ 63
DB 35 VCGVEIVANTDQALGRSLAPHKITLTKGLGAGSKPELGRSAEQOKVDIORMQ 94

QY 64 DSNMLFTIGMGGGTCTGAAPVAVSARELGILTVGVSTPERSEGNPRTLANGVKEL 123
DB 95 GADWVFITAGMGGGTCTGAAPVAVSARELGILTVGVSTPERSEGNPRTLANGVKEL 154

QY 124 AKYVDLIVPNONLALADKSTMLAEFRADVDLLEGVGVYDLIVRPLINL 178
DB 155 KEAVDPLIVIPNDRLQIVDKNTPELEAFREADVNLROGVGIGISDLIVPGLINL 209

RESULT 3
US-10-282-122A-47081
Sequence 47081, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 47081
LENGTH: 404
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-10-282-122A-47081

Query Match 57.2%; Score 506; DB 6; Length 404;
Best Local Similarity 58.7%; Pred. No. 5.5e-39;
Matches 101; Conservative 27; Mismatches 44; Indels 0; Gaps 0;

QY 7 VERIVANTDQALGRSLAPHKITLTKGLGAGSKPELGRSAEQOKVDIORMQ 66
DB 52 VERIVANTDQALGRSLAPHKITLTKGLGAGSKPELGRSAEQOKVDIORMQ 111

QY 67 MLEITGMMGGTCTGAAPVAVSARELGILTVGVSTPERSEGNPRTLANGVKEL 126
DB 112 MLEITGMMGGTCTGAAPVAVSARELGILTVGVSTPERSEGNPRTLANGVKEL 171

QY 127 VDLIVPNONLALADKSTMLAEFRADVDLLEGVGVYDLIVRPLINL 178
DB 155 KEAVDPLIVIPNDRLQIVDKNTPELEAFREADVNLROGVGIGISDLIVPGLINL 209


```
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT FILING DATE: 2003-02-20
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 53331
LENGTH: 394
TYPE: PRT
ORGANISM: Clostridium difficile
US-10-282-122A-53331

Query Match      57.1%; Score 505; DB 6; Length 394;
Best Local Similarity 55.9%; Pred. No. 6; Se-39;
Matches 99; Conservative 35; Mismatches 43; Indels 0; Gaps 0;

OY 2 SOLGEVFIVAMDDCALGRSLAPHKITTKGDKITKGLGAGSKPELGRSABOQKVDIORM 61
DB 42 AQLKGEFIVANTDQALYTSKAEYVQISEKLTGAGANPEVGRKAAESQVISESQ 101
OY 62 LODSNMLFTTGMGGGTCTGAPVAVASVARELGLITVGVVSTPRSEGNPNTRLANAGVK 121
DB 102 LOGADMVEFTAGGGGTGAPVAVAGLAKEMLITGVVYTKPFAFEKIKMKNAEGGIA 161
OY 122 ELAKYVDLIVPNQNLALADSKSTMLEAFRYADVLLBSVKGVTDLIVRGLINL 178
DB 162 ELKSKYVDLITIPNRLQIVOKNTSLDAFAVADVLKQIGISIDLIAVEGLINL 218

RESULT 7
US-10-282-122A-42469
Sequence 42469, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
```

```
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT FILING DATE: 2003-02-20
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 42469
LENGTH: 410
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-10-282-122A-42469

Query Match      55.9%; Score 495; DB 6; Length 410;
Best Local Similarity 57.1%; Pred. No. 6; Se-38;
Matches 100; Conservative 28; Mismatches 47; Indels 0; Gaps 0;

OY 4 LEGVEFIVAMDDCALGRSLAPHKITTKGDKITKGLGAGSKPELGRSABOQKVDIORM 63
DB 36 VKGEVFIVANTDVQALKSKAEYVQIQPKYTRGAGSQPVGQKAAESQVISESQ 95
OY 64 DSNMLFTTGMGGGTCTGAPVAVASVARELGLITVGVVSTPRSEGNPNTRLANAGVKEL 123
DB 96 GADMIFITTAGGGGTGAPVAVAKELGALTGVVYTRPSFEGPKKGRFAEGIAL 155
OY 124 AKYVDLIVPNQNLALADSKSTMLEAFRYADVLLBSVKGVTDLIVRGLINL 178
DB 156 KENVDTLITISNNRLLEVYDKITPMLAEFRADNVLRQGVQGISDITAPGYNL 210

RESULT 8
US-09-134-000C-4658
Sequence 4658, Application US/09134000C
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-002
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4658
LENGTH: 435
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-4658

Query Match      55.9%; Score 495; DB 5; Length 435;
```

Best Local Similarity 57.18; Pred. No. 6,5e-38;
Matches 100; Conservative 28; Mismatches 47; Indels 0; Gaps 0

```

QY 4 LEGVEFIYANDCCOLAGRSIAPHKHTTKGIDIRKGGAGSKPELGKRSABQKXDIOQMLQ 63
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 61 VKGEFTIANDDVQALHKHSKAEFTVYIQLGKRYTRGJAGSQPEVGQAAARESEVOISESI 120
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 64 DSNMLEITGMSGGCTGTGAAPVYASARELGITLVGVSTPFPSEGGPNRTRLNACVKEL 123
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 121 GADMFITRAGMOGGTGTGAAPVYAIKAKELGALITGVYTRPFPSEGGPKGRFAEGTALL 180
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 124 AKYVDFLIVPQNOLLALADKSTMLEAFYADVLEGGKGVTDLIVRGLINTL 178
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 181 KENVDTLLIISNNRLLEVYDKKTPMLEAFREADNVLRGVGGGSDLTITAPGYNLT 235
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

RESULT 9
US-09-134-000C-4658
Sequence 4658

```

: Sequence 4658, Application US/09134000C
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
: TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 032796-032
: CURRENT APPLICATION NUMBER: US/09/134,000C
: CURRENT FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/055,778
: PRIOR FILING DATE: 1997-08-15
: NUMBER OF SEQ ID NOS: 6812
: SOFTWARE: Patentin version 3.1
: SEQ ID NO 4658
: LENGTH: 435
: TYPE: PRN
: ORGANISM: Enterococcus faecalis
: US-09-134-000C-4658

```

Query Match	55.98;	Score 495;	DB 5;	Length 435;
Best Local Similarity	57.18;	Pred. No. 6.5e-38;		
Matches 100; Conservative	28;	Mismatches 47;	Indels 0;	Gaps 0

[illegible]

RESULT 10
US-10-282

Sequence 57592, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Heselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EILTHA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20

```

? PRIOR APPLICATION NUMBER: 60/191,078
? PRIOR FILING DATE: 2000-03-21
? PRIOR APPLICATION NUMBER: 60/206,848
? PRIOR FILING DATE: 2000-05-23
? PRIOR APPLICATION NUMBER: 60/207,727
? PRIOR FILING DATE: 2000-05-26
? PRIOR APPLICATION NUMBER: 60/230,335
? PRIOR FILING DATE: 2000-09-06
? PRIOR APPLICATION NUMBER: 60/230,347
? PRIOR FILING DATE: 2000-09-09
? PRIOR APPLICATION NUMBER: 60/244,578
? PRIOR FILING DATE: 2000-10-23
? PRIOR APPLICATION NUMBER: 60/253,625
? PRIOR FILING DATE: 2000-11-27
? PRIOR APPLICATION NUMBER: 60/257,931
? PRIOR FILING DATE: 2000-12-22
? PRIOR APPLICATION NUMBER: 60/267,636
? PRIOR FILING DATE: 2001-02-09
? PRIOR APPLICATION NUMBER: 60/269,308
? PRIOR FILING DATE: 2001-02-16
? Remaining Prior Application data removed - See File Wrapper or PALM
? NUMBER OF SEQ ID NOS: 78614
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 57592
? LENGTH: 413
? TYPE: PRT
? ORGANISM: Enterococcus faecium
? FEATURE:
? NAME/KEY: MISC FEATURE
? LOCATION: (343)..(343)
? OTHER INFORMATION: x-any amino acid
US-10-262-122A-57592

```

Query Match	55.88;	Score 494;	DB 6;	Length 413;
Best Local Similarity	55.48;	Pred. No. 7.5e-38;		
Matches 97;	Conservative 31;	Mismatches 47;	Indels 0;	Gaps 0

```

QY 4. LEEGVFIYAANDCCQALGSRSLAPHKITLGGDITRGLGAGSKPELGKRSABEQAKYDIQMLQ 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 36 VKGEVFIATANDVQALKNKSKAETVYIQLGKRYTRGLGAGSOPEVGQCAAESEESQIREALD 95
QY 64 DSNMFLITGMCGCCTGAAPYVASAREBLGILTYGVVSTPPESBEPNRRRLNAGVKEL 123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 96 GADMFLFIAGMGCGGTGCAPIYVAGIAKELGALTYGVVTRPFFEEBPKRRFRFAEGIARL 155
QY 124 AKYVDLITVNPONLTLALADKSTMLAEARYDDVLEBVGKGTDLIVRGELNL 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 156 KENDDTLITISNNRLEEVADKFTPMLEAFREADNVLROGVQGISDILTITAGVYNL 210

```

RESULT 11
US-10-417-884-7295

Sequence 7295, Application US/10417884
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: ENTEROCOCCUS AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
City: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10,417,884
FILING DATE: 17-Apr-2003

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/107, 532A
FILING DATE: 30-Jun-1998
APPLICATION NUMBER: 60/085, 598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Atinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 7295:
SEQUENCE CHARACTERISTICS:
LENGTH: 414 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc.feature
LOCATION: (8) LOCATION 1...414
SEQUENCE DESCRIPTION: SEQ ID NO: 7295:
US-10-417-884-7295

Query Match 55.8%; Score 494; DB 6; Length 414;
Best Local Similarity 55.4%; Pred. No. 7.5e-38;
Matches 97; Conservative 31; Mismatches 47; Indels 0; Gaps 0;

QY 4 LKGEFVYANTDQALGSLAPHKITLTKGDKGAGSKPELGRSAEQKVDIQRLQ 63
DB 37 VAGVEFVYANTDQALGSLAPHKITLTKGDKGAGSKPELGRSAEQKVDIQRLQ 96
QY 64 DSNMFLITGMSGGTCTGAPVAVASVARELGILTVGVSTPFRSEGNFTRLANAGVKEL 123
DB 97 GADMFTITGMSGGTCTGAPVAVASVARELGILTVGVSTPFRSEGNFTRLANAGVKEL 156
QY 124 AKYVDTLIVPNQNLALADKSTMTLEAFRYADDVLLGKGVTDLIVRPGILNL 178
DB 157 KENVDTLIIISNNRLELVVYDKTPMLEAFREADNVLRQGVGIDSLITAPGVNL 211

RESULT 12
US-10-282-122A-61968
Sequence 61968, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 61968
LENGTH: 387
TYPE: PR1
ORGANISM: Mycobacterium avium
US-10-282-122A-61968

Query Match 55.6%; Score 492; DB 6; Length 387;
Best Local Similarity 54.9%; Pred. No. 1.1e-37;
Matches 96; Conservative 35; Mismatches 44; Indels 0; Gaps 0;

QY 4 LKGEFVYANTDQALGSLAPHKITLTKGDKGAGSKPELGRSAEQKVDIQRLQ 63
DB 33 LKGEFVYANTDQALGSLAPHKITLTKGDKGAGSKPELGRSAEQKVDIQRLQ 92
QY 64 DSNMFLITGMSGGTCTGAPVAVASVARELGILTVGVSTPFRSEGNFTRLANAGVKEL 123
DB 93 GADMFTITGMSGGTCTGAPVAVASVARELGILTVGVSTPFRSEGNFTRLANAGVKEL 152
QY 124 AKYVDTLIVPNQNLALADKSTMTLEAFRYADDVLLGKGVTDLIVRPGILNL 178
DB 153 RESCDTIVIPNDRLDMGDAVSLMDAFRSADVLNGLVGGITDILITPGILNV 207

RESULT 13
US-10-282-122A-52722
Sequence 52722, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27

```
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52722
; LENGTH: 369
; TYPE: PRF
; ORGANISM: Clostridium botulinum
US-10-282-122A-52722
```

```
Query Match
Best Local Similarity 53.1%; Score 490; DB 1; Length 369;
Matches 93; Conservative 38; Mismatches 44; Indels 0; Gaps 0;
```

```
QY 4 LECVEFIYANTDQALGRSLAPRKITLKGDKITGLGAGSKPELGKRSAGQKVDIORMQ 63
DB 35 LKVEFIYANTDQALGRSLAPRKITLKGDKITGLGAGSKPELGKRSAGQKVDIORMQ 94
QY 64 DSMMLFITGGMGGCTGGAAPVAVASVARELGITLVGVSTPFRSEGNRRLANAGVEL 123
DB 95 GADMVFTAGMGCTGGAAPVAVASVARELGITLVGVSTPFRSEGNRRLANAGVEL 154
QY 124 AKYVDLIVPNONLALADKSTMTLEAFRYADVDLLEGVKGVTDLIVRGLINL 178
DB 155 REHVDTLVITISNRLLEIYDKTTPMMEAFREADVNLQGVGISDLITINPGYVNL 209
```

```
RESULT 14
PCT-US02-36122-84
; Sequence 84, Application PC/TUS0236122
; GENERAL INFORMATION:
; APPLICANT: Murphy, Ellen and Projan, Stephen, J.
; TITLE OF INVENTION: Alloiococcus oltidis Infectious Disease Targets
; FILE REFERENCE: Application 1
; CURRENT APPLICATION NUMBER: PCT/US02/36122
; CURRENT FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 84
; LENGTH: 410
; TYPE: PRF
; ORGANISM: Alloiococcus oltidis
PCT-US02-36122-84
```

```
Query Match
Best Local Similarity 55.4%; Score 490; DB 1; Length 410;
Matches 92; Conservative 37; Mismatches 46; Indels 0; Gaps 0;
```

```
QY 4 LECVEFIYANTDQALGRSLAPRKITLKGDKITGLGAGSKPELGKRSAGQKVDIORMQ 63
DB 36 VEGVEFIYANTDQALGRSLAPRKITLKGDKITGLGAGSKPELGKRSAGQKVDIORMQ 95
QY 64 DSMMLFITGGMGGCTGGAAPVAVASVARELGITLVGVSTPFRSEGNRRLANAGVEL 123
DB 96 GADMVFTAGMGCTGGAAPVAVASVARELGITLVGVSTPFRSEGNRRLANAGVEL 155
QY 124 AKYVDLIVPNONLALADKSTMTLEAFRYADVDLLEGVKGVTDLIVRGLINL 178
DB 156 REHVDTLVITISNRLLEIYDKTTPMMEAFREADVNLQGVGISDLITINPGYVNL 210
```

```
RESULT 15
PCT-US02-36123-5202
; Sequence 5202, Application PC/TUS0236123
; GENERAL INFORMATION:
; APPLICANT: American Cyanamid Company, and Fletcher, Leah D., McMichael, John C.,
; APPLICANT: Russell, David P., and Zagursky, Robert J.
; TITLE OF INVENTION: Alloiococcus oltidis Open Reading Frames (ORFs) Encoding Polypep
```

```
; TITLE OF INVENTION: Antigens, Immunogenic Compositions and Uses Thereof
; FILE REFERENCE: Application 1
; CURRENT APPLICATION NUMBER: PCT/US02/36123
; CURRENT FILING DATE: 2003-01-02
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5202
; LENGTH: 410
; TYPE: PRF
; ORGANISM: Alloiococcus oltidis
PCT-US02-36123-5202
```

```
Query Match
Best Local Similarity 52.6%; Score 490; DB 1; Length 410;
Matches 92; Conservative 37; Mismatches 46; Indels 0; Gaps 0;
```

```
QY 4 LECVEFIYANTDQALGRSLAPRKITLKGDKITGLGAGSKPELGKRSAGQKVDIORMQ 63
DB 36 VEGVEFIYANTDQALGRSLAPRKITLKGDKITGLGAGSKPELGKRSAGQKVDIORMQ 95
QY 64 DSMMLFITGGMGGCTGGAAPVAVASVARELGITLVGVSTPFRSEGNRRLANAGVEL 123
DB 96 GADMVFTAGMGCTGGAAPVAVASVARELGITLVGVSTPFRSEGNRRLANAGVEL 155
QY 124 AKYVDLIVPNONLALADKSTMTLEAFRYADVDLLEGVKGVTDLIVRGLINL 178
DB 156 REHVDTLVITISNRLLEIYDKTTPMMEAFREADVNLQGVGISDLITINPGYVNL 210
```

```
Search completed: June 2, 2003, 08:48:40
Job time : 42 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2003, 07:18:30 ; Search time 40 Seconds
(without alignments)
427.798 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASOLEGEFIVANTDCOALG.....LLEGVKGVTLIVRGLINL 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	524	59.2	452	2	G97826 cell division prot
2	518	58.5	452	2	F71672 cell division prot
3	511	57.7	382	2	P83969 cell division prot
4	510	57.6	391	2	AH1328 cell division prot
5	510	57.6	392	2	AH1699 cell division prot
6	510	57.6	398	2	S35264 cell division prot
7	507	57.3	373	2	H97108 cell division prot
8	505	57.2	404	2	C70137 cell division prot
9	505	57.1	508	2	C87564 cell division prot
10	498	56.3	382	1	I39848 cell division prot
11	493	55.7	417	2	E86858 cell division prot
12	486	54.9	399	1	S60765 cell division prot
13	485	54.8	379	2	JEO282 cell division prot
14	485	54.8	379	2	B70579 cell division prot
15	482	54.5	583	2	C97610 cell division prot
16	481	54.4	373	2	AG2832 cell division prot
17	481	54.4	373	2	JC7087 cell division prot
18	480	54.4	379	2	G87023 cell division prot
19	480	54.2	390	1	S58814 cell division prot
20	475	53.8	538	2	AC3325 cell division prot
21	475	53.7	590	1	A38119 cell division prot
22	472	53.3	419	2	H95193 cell division prot
23	472	53.3	419	2	E98060 cell division prot
24	471	53.2	400	2	T10476 cell division prot
25	468	52.9	379	1	JC4289 cell division prot
26	468	52.9	428	2	AC2288 cell division prot
27	467	52.8	428	4	S46932 cell division prot
28	467	52.8	413	2	B82760 cell division prot
29	467	52.8	430	2	S77393 cell division prot

30	465	52.5	418	2	H71331 probable cell divi
31	464	52.4	394	2	H83093 cell division prot
32	464	52.3	371	2	E75494 cell division prot
33	463	52.3	397	2	E84778 plastid division p
34	463	52.3	478	2	JC7770 chloroplast divisi
35	462	52.2	464	2	T51090 plastid division p
36	461	52.1	384	2	A84955 plastid division p
37	461	52.1	458	2	T51089 plastid division p
38	460.5	52.0	371	2	T51695 cell division prot
39	460	52.0	383	2	AE0069 cell division prot
40	459	51.9	345	1	S58854 cell division prot
41	458	51.8	473	2	T49028 plastid division p
42	455	51.4	399	2	G90102 cell division prot
43	453	51.2	383	2	AB0519 cell division prot
44	452	51.1	383	1	CEECZ cell division prot
45	452	51.1	383	2	C90641 cell division prot

ALIGNMENTS

RESULT 1
G97826
cell division protein ftsz [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C:Date: 30-Sep-2001 #sequence_rev10 30-Sep-2001 #text_change 17-May-2002
C:Accession: G97826
R: Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: G97826
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-452 <NOR>
A:Cross-references: GB:AE006914; PIDN:AL03553.1; PID:915620131; GSPDB:GN00173
C:Genetics:
A:Gene: ftsz
C:Superfamily: cell division protein ftsz

Query Match 59.28; Score 524; DB 2; Length 452;
Best Local Similarity 56.28; Pred. No. 2.2e-38;
Matches 100; Conservative 38; Mismatches 40; Indels 0; Gaps 0;

QY 1 ASOLEGEFIVANTDCOALGSLAPHKITTKGDKTGAGSKPELGRSAEOKVDIOR 60
DB 35 SANLQGANFVANTDAOGLSHSLCTNKIQGVSTTRGAGASPEVGALAOSESSEIRN 94
QY 61 MLODSNMLFITGGMGCGCTGAAPVAVASVARELGIIVGVVSTPFRSEGPNTRLANGV 120
DB 95 YLENSMNVFTTAGMGCGTGSAPVIRAKELGIIIVGVYTKRPFHFGCHRMKADAGL 154
QY 121 KELAKYVDITLIVPNQNLALADKSTMLEAFRYADVDLLEGVKGVTLIVRGLINL 178
DB 155 IELQGEFVDLIVPNQNLAFRIANEOTTFADAFKADVDLHAGVGVTLIMPGLINL 212

RESULT 2
F71672
cell division protein ftsz (fts2) RP666 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_rev10 21-Nov-1998 #text_change 03-Nov-2000
C:Accession: F71672
R: Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Almark
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499; PMID:9823893
A:Accession: F71672
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-452 <AND>
A:Cross-references: GB:AJ235272; GB:AJ235269; NID:93861033; PIDN:CAA15104.1; PID:6134
A:Experimental source: strain Madrid E

C:Genetics:
A:Gene: ftsZ; RP666
C:Superfamily: cell division protein ftsZ

Query Match 58.5%; Score 518; DB 2; Length 452;
Best Local Similarity 55.9%; Pred. No. 7.2e-38;
Matches 99; Conservative 38; Mismatches 40; Indels 0; Gaps 0;

OY 2 SOLEGEFIVAMTDCOALGRSLAPHKITTKGDKITGAGSKPELGRSAEQKVDIORM 61
DB 36 AMIQGANFVAMTDCOALGRSLAPHKITTKGDKITGAGSKPELGRSAEQKVDIORM 95
OY 62 LODSNMFLITGGMGCTGGAAPVAVASVARELGILTVGVSTPRSEGNPRTRLANAGVK 121
DB 96 LENSMMVFITAGMGCTGGAAPVAVASVARELGILTVGVSTPRSEGNPRTRLANAGVK 155
OY 122 ELAKYVDLIVPNQNLALADKSTMTLEAFRYADVLEBVGKVTDLIVRGLINTL 178
DB 156 ELQGEFVDTLIVPNQNLALADKSTMTLEAFRYADVLEBVGKVTDLIVRGLINTL 212

RESULT 3

F83969

cell-division initiation protein (septum formation) ftsZ [imported] - Bacillus halodurans
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: F83969

R.Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: AB3650; MUID:20512582; PMID:11058132

A:Accession: F83969

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-382 <STO>

A:Cross-references: GB:AP001515; GB:BA000004; NID:910174886; PIDN:BA06277.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: ftsZ

C:Superfamily: cell division protein ftsZ

Query Match 57.7%; Score 511; DB 2; Length 382;
Best Local Similarity 58.3%; Pred. No. 2.5e-37;
Matches 102; Conservative 30; Mismatches 43; Indels 0; Gaps 0;

OY 4 LEGVEFIVANTDCOALGRSLAPHKITTKGDKITGAGSKPELGRSAEQKVDIORM 63
DB 35 LOGVDFISVNTDAQALHLKAEVKLQCGKLTGAGANPEIGKKAESREQIEALQ 94
OY 64 DSNMFLITGGMGCTGGAAPVAVASVARELGILTVGVSTPRSEGNPRTRLANAGVK 123
DB 95 GADMFVITAGMGCTGGAAPVAVASVARELGILTVGVSTPRSEGNPRTRLANAGVK 154
OY 124 AKYVDLIVPNQNLALADKSTMTLEAFRYADVLEBVGKVTDLIVRGLINTL 178
DB 155 KEAVDTLIVPNQNLALADKSTMTLEAFRYADVLEBVGKVTDLIVRGLINTL 209

RESULT 4

AH1328

cell-division initiation protein ftsZ homolog ftsZ [imported] - Listeria monocytogenes
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AH1328

R.Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

Science 294, 849-852, 2001

A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schlueter, T.; Simoes, N.; Tlerez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant,

A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1328

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-391 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00110.1; PID:916411502; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: ftsZ
C:Superfamily: cell division protein ftsZ

Query Match 57.6%; Score 510; DB 2; Length 391;
Best Local Similarity 56.0%; Pred. No. 3.1e-37;
Matches 98; Conservative 35; Mismatches 42; Indels 0; Gaps 0;

OY 4 LEGVEFIVAMTDCOALGRSLAPHKITTKGDKITGAGSKPELGRSAEQKVDIORM 63
DB 35 VQGEFISVNTDAQALHLKAEVKLQCGKLTGAGANPEIGKKAESREQIEALQ 94
OY 64 DSNMFLITGGMGCTGGAAPVAVASVARELGILTVGVSTPRSEGNPRTRLANAGVK 123
DB 95 GSDMFVITAGMGCTGGAAPVAVASVARELGILTVGVSTPRSEGNPRTRLANAGVK 154
OY 124 AKYVDLIVPNQNLALADKSTMTLEAFRYADVLEBVGKVTDLIVRGLINTL 178
DB 155 KEAVDTLIVPNQNLALADKSTMTLEAFRYADVLEBVGKVTDLIVRGLINTL 209

RESULT 5

AH1699

cell-division initiation protein ftsZ homolog ftsZ [imported] - Listeria innocua (str

C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AH1699

R.Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,

Science 294, 849-852, 2001

A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.;

ok, C.; Schlueter, T.; Simoes, N.; Tlerez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1699

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-392 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC97368.1; PID:916414652; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: ftsZ

C:Superfamily: cell division protein ftsZ

Query Match 57.6%; Score 510; DB 2; Length 392;
Best Local Similarity 56.0%; Pred. No. 3.1e-37;
Matches 98; Conservative 35; Mismatches 42; Indels 0; Gaps 0;

OY 4 LEGVEFIVAMTDCOALGRSLAPHKITTKGDKITGAGSKPELGRSAEQKVDIORM 63
DB 35 VQGEFISVNTDAQALHLKAEVKLQCGKLTGAGANPEIGKKAESREQIEALQ 94
OY 64 DSNMFLITGGMGCTGGAAPVAVASVARELGILTVGVSTPRSEGNPRTRLANAGVK 123
DB 95 GSDMFVITAGMGCTGGAAPVAVASVARELGILTVGVSTPRSEGNPRTRLANAGVK 154
OY 124 AKYVDLIVPNQNLALADKSTMTLEAFRYADVLEBVGKVTDLIVRGLINTL 178
DB 155 KEAVDTLIVPNQNLALADKSTMTLEAFRYADVLEBVGKVTDLIVRGLINTL 209

RESULT 6

S35264

cell division protein ftsZ homolog - Wolbachia sp. (fregment)

C:Species: Wolbachia sp.
C:Date: 09-Dec-1993 #sequence_revision 12-Apr-1996 #text_change 26-Aug-1999
C:Accession: S35264; S35404

R:Holden, P.R.; Brookfield, J.F.Y.; Jones, P.

Mol. Gen. Genet. 240, 213-220, 1993

A:Title: Cloning and characterization of an *ftsZ* homologue from a bacterial symbiont of

A:Reference number: S35264; MUID:93360900; PMID:7689140

A:Accession: S35264

A:Molecule type: DNA

A:Residues: 1398 <HOL>

A:Cross-references: EMBL:X71906; NID:9311274; PIDN:CAA50724.1; PID:9311275

A:Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in lacking 284-G)

C:Genetics:

A:Gene: *ftsZ*

A:Cross-references: Flybase:Flybase013356

C:Superfamily: cell division protein *ftsZ*

C:Keywords: cell division; GTP binding

F:109-115/Region: tubulin/*ftsZ* GTP/GDP-binding (G-G-G-T-G-[ST]-G) motif

Query Match 57.6%; Score 510; DB 2; Length 398;

Best Local Similarity 56.1%; Pred. No. 3.1e-37; Matches 106; Conservative 29; Mismatches 42; Indels 12; Gaps 1;

OY 2 SOLEGEFIVANTDQALGRSLAPHKITLTKGAGSKPELGRSAEQKVDIORM 61

Db 36 SNIQGVNFVANNDQALGRSLAPHKITLTKGAGSKPELGRSAEQKVDIORM 95

OY 62 LODSNMLFTGGMGCTGTAAPVAVASVAREL-----GILTVGVSTPFRSEG 109

Db 96 IKDSHMLFTAGMGCGTGTGAAPVAVAKARARAAVADRAPKEKKILTVGVSTPFRSEG 155

OY 110 PNRTRLANGVRLKRLAYVDLIVPNONLALADKSTTMEAFKADVDLLEKGVYDL 169

Db 156 VRMPRAEGLLELQKRYVDLIVPNONLFRANKEKTFSDAFKADVDLHIGIRYDL 215

OY 170 IVRGLINL 178

Db 216 MWMPGLINL 224

RESULT 7

H97108 cell division GTPase *ftsZ* [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001

C:Accession: H97108

R:Noelling, J.; Breton, G.; Omelchenko, M.V.; Markova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 193, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C1C

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: H97108

A:Molecule type: DNA

A:Status: preliminary

A:Residues: 1-373 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK79659.1; PID:915024657; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC1693

C:Superfamily: cell division protein *ftsZ*

Query Match 57.3%; Score 507; DB 2; Length 373;

Best Local Similarity 56.0%; Pred. No. 5.4e-37; Matches 98; Conservative 35; Mismatches 42; Indels 0; Gaps 0;

OY 4 LEGVEPIVANTDQALGRSLAPHKITLTKGAGSKPELGRSAEQKVDIORM 63

Db 35 LKNVEPIVANTDQALGRSLAPHKITLTKGAGSKPELGRSAEQKVDIORM 94

OY 64 DSMLEFITGGMGCTGTAAPVAVASVARELGLTVGVSTPFRSEGNRRLANAGVEL 123

Db 95 GADWEITAGMGCGTGTGAAPVAVAEIAKSMGILTVGVSTPFRSEGNRRLANAGVEL 154

OY 124 AKYVDLIVPNONLALADKSTTMEAFKADVDLLEKGVYDLIVRGLINL 178

Db 155 KENVDTLIVPNONLALADKSTTMEAFKADVDLLEKGVYDLIVRGLINL 209

RESULT 8

C70137

cell division protein *ftsZ* - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 26-Aug-1999

C:Accession: C70137; I40082; S21557

R:Fraser, C.M.; Castens, S.; Huang, M.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh

son, D.; Peterson, J.; Keriavage, A.R.; Quackenbush, J.; Salberg, S.; Hanson, M.; Yu

; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: C70137

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-404 <LE>

A:Cross-references: GB:AE001137; GB:AE000783; NID:92688160; PIDN:AAC6649.1; PID:9268

A:Experimental source: strain B31

R:Old, I.G.; MacDougall, J.H.; Saint Girons, I.; Davidson, B.E.

FEBS Microbiol. Lett. 99, 245-250, 1992

A:Title: Mapping of genes on the linear chromosome of the bacterium Borrelia burgdorferi

A:Reference number: I40082

A:Accession: I40082

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 34-119, 'A' <RES>

A:Cross-references: EMBL:212164; NID:949293; PID:949294

A:Experimental source: strain 212

C:Comment: This cytoplasmic protein associates with the cell membrane just prior to c

C:Genetics:

A:Gene: *ftsZ*

C:Superfamily: cell division protein *ftsZ*

C:Keywords: cell division; GTP binding

F:120-126/Region: tubulin/*ftsZ* GTP/GDP-binding (G-G-G-T-G-[ST]-G) motif

F:221-228/Region: GTP-hydrolyzing region

Query Match

Best Local Similarity 58.7%; Score 506; DB 2; Length 404;

Matches 101; Conservative 27; Mismatches 44; Indels 0; Gaps 0;

OY 7 VEFIVANTDQALGRSLAPHKITLTKGAGSKPELGRSAEQKVDIORM 66

Db 52 VEFIVANTDQALGRSLAPHKITLTKGAGSKPELGRSAEQKVDIORM 111

OY 67 MEFITGGMGCTGTAAPVAVASVARELGLTVGVSTPFRSEGNRRLANAGVEL 126

Db 112 MEFITGGMGCTGTAAPVAVASVARELGLTVGVSTPFRSEGNRRLANAGVEL 171

OY 127 VDTLIVPNONLALADKSTTMEAFKADVDLLEKGVYDLIVRGLINL 178

Db 172 VDTLIVPNONLALADKSTTMEAFKADVDLLEKGVYDLIVRGLINL 223

RESULT 9

C87564 cell division protein *ftsZ* [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 17-May-2002

C:Accession: C87564

R:Nlema, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko

n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: C87564

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-508 <STO>

A:Cross-references: GB:AE005673; NID:913424103; PIDN:AAK24511.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC2540

C:Superfamily: cell division protein ftsz1

Query Match 57.1%; Score 505; DB 2; Length 508;

Best Local Similarity 58.3%; Pred. No. 1.2e-36;

Matches 102; Conservative 26; Mismatches 45; Indels 0; Gaps 0;

QY 4 LEGVEFIVANTDQALGRSLAPHKITLTKGLGAGSKPELGRSAEQOKVDIORMLO 63

DB 38 LEGVEFIVANTDQALQLOFAKTRDRIRQLGVQITGAGAHPEVGSAAEESPEIGEHLD 97

QY 64 DSNMFLITGMSGGCTCGAPVAVASVARELGILTVGVSTPFRSEGNPRTLANAGVKEL 123

DB 98 GAHMVFITGMSGGCTGAPVAVASVARELGILTVGVSTPFRSEGNPRTLANAGVKEL 157

QY 124 AKYVDLIVPNQNLALADKSTTMEAFRRADVDLLEGVGVTDLIVRGLINL 178

DB 158 QRYVDLIVPNQNLFRVANERTFAPAFGMADVLSHGSVSTIDMLVPLGLINL 212

RESULT 10

139848 cell division initiation protein (septum formation) Ftsz - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 19-Jul-1996 #sequence_revision 04-Oct-1996 #text_change 16-Jun-2000

C:Accession: I39848; A69628

R:Beall, B.; Lowe, M.; Lutkenhaus, J.

J. Bacteriol. 170, 4855-4864, 1988

A:Title: Cloning and characterization of Bacillus subtilis homologs of Escherichia coli

A:Reference number: I39846; MUID:89008108; PMID:3139638

A:Accession: I39848

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-382 <RES>

A:Cross-references: GB:M22630; NID:q142938; PIDN:AAA22457.1; PID:q142941

R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beret

C.: Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Cho

A.: Ehrlich, S.D.; Emerson, P.T.; Ertian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 380, 249-256, 1997

A:Authors: Foulger, D.; Fultz, C.; Fujita, M.; Fujita, Y.; Funo, S.; Galizzi, A.; Gall

leach, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

Y. M.; Ogawa, K.; Ogawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadleir, Y.; Sato, T.; Scanlon,

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot

A.: Winters, P.; Wipat, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: A69628

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-382 <KUN>

A:Cross-references: GB:Z99111; GB:AL009126; NID:92633699; PIDN:CAB13402.1; PID:92633900

A:Experimental source: strain 168

C:Comment: This cytoplasmic protein associates with the cell membrane just prior to cell

ch vegetative and sporulating cell divisions.

C:Genetics:

A:Gene: ftsz

C:Superfamily: cell division protein ftsz

C:Keywords: cell division; GTP binding

F106-112/Region: tubulin/fts2 GTP-binding (G-G-G-T-G-[ST]-G) motif

Query Match 56.3%; Score 498; DB 1; Length 382;

Best Local Similarity 53.7%; Pred. No. 3.4e-36;

Matches 95; Conservative 37; Mismatches 45; Indels 0; Gaps 0;

QY 2 SOLEGEFIVANTDQALGRSLAPHKITLTKGLGAGSKPELGRSAEQOKVDIORM 61

DB 33 NEVGVEFIVANTDQALGRSLAPHKITLTKGLGAGSKPELGRSAEQOKVDIORM 92

QY 62 LQDSNMLEITGMSGGCTCGAPVAVASVARELGILTVGVSTPFRSEGNPRTLANAGVK 121

DB 93 LKADAMFVTAAGMGCTGAPVAVASVARELGILTVGVSTPFRSEGNPRTLANAGVK 152

QY 122 ELAKYVDLIVPNQNLALADKSTTMEAFRRADVDLLEGVGVTDLIVRGLINL 178

DB 153 AMKEAVDVLIVPNDRILEIVDKNPMLEAFREADVLSHGSVSTIDMLVPLGLINL 209

RESULT 11

E86858 cell division protein Ftsz [imported] - Lactococcus lactis subsp. lactis (strain IL14

C:Species: Lactococcus lactis subsp. lactis

C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: E86858

R:Boletín, A.; Wincker, P.; Mauger, S.; Jallón, O.; Malarne, K.; Weissenbach, J.; Eh

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis

A:Reference number: A86825; MUID:21235186; PMID:11337471

A:Accession: E86858

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-417 <SMO>

A:Cross-references: GB:AB005176; PID:q12724901; PIDN:AAK05967.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: ftsz

C:Superfamily: cell division protein ftsz

Query Match 55.7%; Score 493; DB 2; Length 417;

Best Local Similarity 56.0%; Pred. No. 1e-35;

Matches 98; Conservative 29; Mismatches 48; Indels 0; Gaps 0;

QY 4 LEGVEFIVANTDQALGRSLAPHKITLTKGLGAGSKPELGRSAEQOKVDIORMLO 63

DB 36 VSGVEFIVANTDQALGRSLAPHKITLTKGLGAGSKPELGRSAEQOKVDIORMLO 95

QY 64 DSNMFLITGMSGGCTCGAPVAVASVARELGILTVGVSTPFRSEGNPRTLANAGVKEL 123

DB 96 GSDMIFITGMSGGCTCGAPVAVASVARELGILTVGVSTPFRSEGNPRTLANAGVKEL 155

QY 124 AKYVDLIVPNQNLALADKSTTMEAFRRADVDLLEGVGVTDLIVRGLINL 178

DB 156 RANVDLILISNNLLEIVDKRTPLEALREADNVLRGVGVTDLITNPGLINL 210

RESULT 12

S60765 cell division protein ftsz - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 27-Apr-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000

C:Accession: S60765; T34952

R:McCormick, J.R.; Su, E.P.; Dricks, A.; Losick, R.

Mol. Microbiol. 14, 243-254, 1994

A:Title: Growth and viability of Streptomyces coelicolor mutant for the cell division

A:Reference number: S60763; MUID:95131746; PMID:7830569

A:Accession: S60765

A:Molecule type: DNA

A:Residues: 1-399 <MCC>

A:Cross-references: EMBL:U0879; NID:94204101; PIDN:AA010533.1; PID:9527649

A:Experimental source: strain A3(2)

R:Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Randsdram, M

submitted to the EMBL Data Library, August 1999

A:Reference number: Z21563

A:Accession: T34952

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-399 <SAU>

A:Cross-references: EMBL:AL109663; PIDN:CAB51991.1; GSPDB:GN00070; SCOEDB:fts2

A:Experimental source: strain A3(2)

C:Comment: This cytoplasmic protein associates with the cell membrane just prior to c

A:Gene: *ftsZ*
A:Start codon: GTC
C:Superfamily: cell division protein *ftsZ*
C:Keywords: cell division: GTP binding
P:103-109/Region: tubulin/*ftsZ* GTP/GDP-binding (G-G-G-I-G-[ST]-G) motif
P:204-211/Region: GTP hydrolyzing
P:383-391/Region: hydrophobic

Query Match	54.9%	Score 466	DB 1	Length 399
Best Local Similarity	53.1%	Pred. No.	4.1e-35	
Matches 93	Conservative 38	Mismatches 44	Indels 0	Gaps 0

[illegible]

RESULT 13
JE0282

C: *Streptomyces collinus*
C: Species: *Streptomyces collinus*
C: Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 18-Aug-2000
C: Accession: J0282
R: Zhulanova, E.; Mikulik, K.
B: Blochem. Biophys. Res. Commun. 249, 556-561, 1998
A: Title: Characterization of *ftsZ* gene and its protein product from *Streptomyces*
#: Reference number: J0282; MUID:95380301; PMID:9712736

A:Accession: J60282
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-402 <ZNU>
A:Cross-references: GB:AF073487; NID:g3328123; PID:AC3005.1; PID:g3328124
C:Genes:
A:Gene: fls2
C:Function:
A:Description: required for septation and conversion of aerial mycelium into ch
C:Superfamily: cell division protein fls2
;Keywords: cell division; GTP binding

Query Match	54.9%	Score 486;	DB 2;	Length 402;
Best Local Similarity	53.1%	Pred. No. 4.1e-35;		
Matches	93;	Conservative	38;	Mismatches 44; Indels 0; Gaps 0;
Qy	4	LEGEVFIVAMDDCALGRSLAPHRKITGKDDITGKGLGAGSPDELGRSAEQGQKVDIQRMQ	63	
	: : :	: : :	: : :	: : :
Db	32	LKGEVFIAINTDADALLMSDADVDLDVGLRELTRGLGAGANPAVGRKAAEDHNEELTEYVK	91	
Qy	64	DSNNLFITGGGGGCTGAPVAVASVARELILTVGVVSTPFRSBRGPARTLRANAGVEL	123	
	::: :	: : :	: : :	: : :
Db	92	GADNVFTYAGGGGTGTGGAPVAVANIASLTALTITGVTRPTFTGRRRANOEDGIAEL	151	
Qy	124	AKYVDTLIVPNQULLALADKSTMLEAFRAADVDLLEGVGVMDLIIRPGILN	178	
	: : :	: : :	: : :	: : :
Db	152	REEVDTLIVPNDRLLISDRQVSLAOFKASADVDLGGVGITDLITTPGLIN	206	

RESULT 14

B/03/79
Probable cell division protein FtsZ - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: B/03/79
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.;
Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; H.
; Alandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998
A:Authors: Squires, R., Sulston, J.E., Taylor, K., Whitehead, S., Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70579

A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <COL>
A:Cross-references: GB:295388; GB:AL123456; NID:93261759; PIDN:CAE08643.1; PID:92104
A:Experimental source: strain H37Rv
A:Genetics:
C:Gene: ftsZ
C:GeneFamily: cell division protein ftsZ

[illegible]

RESULT 15
C97610

Cell division protein ftsZ [imported] - Agrobacterium tumefaciens (strain C58, Cerec
C.Species: Agrobacterium tumefaciens
C.Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C.Accession: C57610
R.Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Gold
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Marelitz,
Science 294, 2323-2328, 2001
A.Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A.Reference number: A97359; PMID:11743194

A:Accession: C97610
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-583 <KUR>
C:Cross-references: GB:AE007869; PIDN:AAK67836.1; PID:g15157218; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_3784
A:Map position: circular chromosome
C:Superfamily: cell division protein ftsZ1

Query Match	54.5%	Score 482	DB 2	Length 583
Best Local Similarity	57.1%	Pred. No. 1,4e-34		
Matches 100	Conservative 26	Mismatches 49	Indels 0	Gaps 0
QY	4	LEGEVEFIATNDCOALGRSLAPNPKITTLCKDITGLGAGSGPELGRSAEQKVDIORMI	63	
		: : :		: : :
Db	38	LQGDVFAVANTDAQILMTKADRVITQLVGNVTBEGAGSGPEVGRRAAECEIDEITDHLN	97	
QY	64	DSNNLFITGGMGCGTGAAPVAVASARELGILTVGVSTPSPRESEGNPNTRLANAGVKEL	123	
		::: : : :		::: : : :
Db	98	GTHMCFYTAGMGCGGTGAAPVVAOARRKKGLITGVVTKPPHFEGRMRRLAEQCIIEEL	157	
QY	124	AKYVDTLIVFNQNLALADKSTYMLAEFRYADVDLLEGVKSTVDLIVRPLGIL	178	
		: : : : : : :		: : : : : : :
Db	158	QKSVDTLIVFNQNLFRILANDKITTFADFAAMADQVLYSGVACITDLMKVEGLINL	212	

Search completed: June 2, 2003, 08:42:15
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 2, 2003, 06:02:36 ; Search time 22 Seconds
(Without alignments)
335.581 Million cell updates/sec

Title: US-09-770-509-2

Perfect score: 885

Sequence: 1 ASQLEGEFIVANTDQALG.....LLEGVKGVTDLVPRGLINL 178

Scoring table: BLOSUM62

Gapop 10.0 , Capext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt.40.*

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	534	59.2	452	1	FTSZ_RICCN
2	518	58.5	452	1	FTSZ_RICCN
3	511	57.7	382	1	FTSZ_BACHD
4	510	57.6	398	1	FTSZ_WOISP
5	506	57.2	399	1	FTSZ_BORBU
6	505	57.1	508	1	FTSZ_CACRC
7	498	56.3	382	1	FTSZ_BACSU
8	494	55.8	392	1	FTSZ_BACSU
9	492	55.6	413	1	FTSZ_BACSU
10	489	55.3	412	1	FTSZ_BACSU
11	486	54.9	386	1	FTSZ_BACSU
12	485	54.8	379	1	FTSZ_BACSU
13	485	54.8	379	1	FTSZ_BACSU
14	485	54.8	407	1	FTSZ_BACSU
15	482	54.5	383	1	FTSZ_BACSU
16	481	54.4	379	1	FTSZ_BACSU
17	480	54.2	390	1	FTSZ_BACSU
18	475	53.7	381	1	FTSZ_BACSU
19	475	53.7	381	1	FTSZ_BACSU
20	471	53.2	400	1	FTSZ_BACSU
21	470	53.1	442	1	FTSZ_BACSU
22	468	52.8	428	1	FTSZ_BACSU
23	467	52.8	430	1	FTSZ_BACSU
24	465	52.5	384	1	FTSZ_BACSU
25	465	52.5	418	1	FTSZ_BACSU
26	464	52.4	394	1	FTSZ_BACSU
27	461	52.1	384	1	FTSZ_BACSU
28	459	51.9	346	1	FTSZ_BACSU
29	458	51.8	170	1	FTSZ_BACSU
30	453	51.2	384	1	FTSZ_BACSU
31	453	51.2	394	1	FTSZ_BACSU
32	452	51.1	383	1	FTSZ_BACSU
33	450	50.8	386	1	FTSZ_BACSU

ALIGNMENTS

RESULT 1
ID FTSZ_RICCN STANDARD: PRT; 452 AA.
AC 092GV7:
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Cell division protein ftz.
GN FTSZ OR RC1015.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiidae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barde V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii";
RT Science 293:2093-2098(2001).
CC -!- FUNCTION: This protein is essential to the cell division process.
CC it seems to assemble into a dynamic ring on the inner surface of
CC the cytoplasmic membrane at the place where division will occur,
CC and the formation of the ring is the signal for septation to
CC begin. Binds to and hydrolyzes GTP (By similarity).
CC -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic Assemblies at the inner surface
CC of the cytoplasmic membrane (By similarity).
CC -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb.slb.ch).
CC
CC EMBL: AE008653; AL03553.1; -.
CC InterPro: IPR000158; FTsz.
CC InterPro: IPR003008; Tubulin_Ftsz.
CC Pfam: PF00091; tubulin; 1.
CC TIGRFAMs: TIGR00065; ftz; 1.
CC DR PROSITE: PS01134; FTsz_1; 1.
CC DR PROSITE: PS01135; FTsz_2; 1.
CC Cell division; Septation; GTP-binding; Complete proteome.
CC NP_BIND 107 115 GTP (POTENTIAL).
CC FT
CC SQ SEQUENCE 452 AA; 48400 MW; B326101A609A43EA CRC64;
Query Match 59.2%; Score 524; DB 1; Length 452;
Best Local Similarity 56.2%; Pred. No. 1.8e-37;
Matches 100; Conservative 38; Mismatches 40; Indels 0; Gaps 0;
QY 1 ASQLEGEFIVANTDQALGRLAPHKITLKGDKTGLGAGSRPELGRSAGAEQKVDIR 60

```

Db      35 SANLQGANVAVNTDQASLEHSLCTKNKIQGVSTTRGIGAGASPEVGAALAOESESEITN 94
OY      61 MLODSNMLFITGGMGGTCTGAPYVASARELIGLVGVSTPFRSGPRTLANAGV 120
Db      95 YLENSNMVFTTGMGGTGTGAPYVARIKKEGILTVGVVTKPPHFGGHRMKTADKGL 154
OY      121 RELAKYVDTLLVPPNOMLLALADKSTTLMLEAFRYADVLLLEGVGVDTLLVPGILNL 178
Db      155 ELQGFVDTLLVIPPONLFRIANEGTTFADAFKADVDVLAHGVGYDMLMIPGLINTL 212

RESULT 2
FTSZ_RICPR STANDARD; PRT; 452 AA.
AC 09ZCO3;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cell division protein ftz.
GN FTZ OR RP666.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
  Rickettsiaceae; Rickettsiidae; Rickettsia.
OX NCBI_Taxid=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
  Sichertz-Ponten T., Alsmark U.C.W., Pedowski R.M., Naeslund A.K.,
  Eriksson A.-S., Winkler H.H., Kurland C.G.;
  "The genome sequence of Rickettsia prowazekii and the origin of
  mitochondria."
RT Nature 396:133-140(1998).
RL -1- FUNCTION: This protein is essential to the cell-division process.
  It seems to assemble into a dynamic ring on the inner surface of
  the cytoplasmic membrane at the place where division will occur,
  and the formation of the ring is the signal for septation to
  begin. Binds to and hydrolyzes GTP (By similarity).
CC -1- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
  of the cytoplasmic membrane (By similarity).
CC -1- SIMILARITY: BELONGS TO THE FTZ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ235272; CAAL5104.1; -.
DR HSSP: Q57816; 1FSZ..
DR InterPro: IPR000158; Tubulin_FtsZ.
DR InterPro: IPR003008; Tubulin_FtsZ.
DR Pfam: PF00091; tubulin; 1.
DR PRINTS: PR00423; CELLDVIFTSZ.
DR TIGRPFAMS: TIGR00065; ftz. 1.
DR PROSITE: PS01134; FTSZ_1; 1.
DR PROSITE: PS01135; FTSZ_2; 1.
DR Cell division; Septation; GTP-binding; Complete proteome.
FT NP_BIND 107 115 GTP (POTENTIAL).
SQ SEQUENCE 452 AA; 48806 MW; D6AC69F28B3343AE CRC64;

Query Match 58.5%; Score 518; DB 1; Length 452;
Best Local Similarity 55.9%; Pred. No. 5.8e-37;
Matches 99; Conservative 38; Mismatches 40; Indels 0; Gaps 0;

OY 2 SOLEGEVIVANTDQALGRSLAPRKITLTKGIGAGSKPELGRSAEQKVDIORM 61
Db 36 ANLQGANVAVNTDQASLEHSLCTKNKIQGVSTTRGIGAGASPEVGAALAOESESEITN 95

```

```

OY      62 LODSNMLFITGGMGGTCTGAPYVASARELIGLVGVSTPFRSGPRTLANAGV 121
Db      96 LENSNMVFTTGMGGTGTGAPYVARIKKEGILTVGVVTKPPHFGGHRMKTADKGL 155
OY      122 ELAKYVDTLLVPPNOMLLALADKSTTLMLEAFRYADVLLLEGVGVDTLLVPGILNL 178
Db      156 ELQGFVDTLLVIPPONLFRIANEGTTFADAFKADVDVLAHGVGYDMLMIPGLINTL 212

RESULT 3
FTSZ_BACHD STANDARD; PRT; 382 AA.
AC 09K97;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cell division protein ftz.
GN FTZ OR BH258.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=66665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
  Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
  Horikoshi K.;
  "Complete genome sequence of the alkaliphilic bacterium Bacillus
  halodurans and genomic sequence comparison with Bacillus subtilis."
RT Nucleic Acids Res. 28:4317-4331(2000).
RL -1- FUNCTION: This protein is essential to the cell-division process.
  It seems to assemble into a dynamic ring on the inner surface of
  the cytoplasmic membrane at the place where division will occur,
  and the formation of the ring is the signal for septation to
  begin. Binds to and hydrolyzes GTP (By similarity).
CC -1- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
  of the cytoplasmic membrane (By similarity).
CC -1- SIMILARITY: BELONGS TO THE FTZ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP001515; BAB06277.1; -.
DR HSSP: Q57816; 1FSZ..
DR InterPro: IPR000158; FtsZ.
DR InterPro: IPR003008; Tubulin_FtsZ.
DR Pfam: PF00091; tubulin; 1.
DR PRINTS: PR00423; CELLDVIFTSZ.
DR TIGRPFAMS: TIGR00065; ftz. 1.
DR PROSITE: PS01134; FTSZ_1; 1.
DR PROSITE: PS01135; FTSZ_2; 1.
DR Cell division; Septation; GTP-binding; Complete proteome.
FT NP_BIND 104 112 GTP (POTENTIAL).
SQ SEQUENCE 382 AA; 40583 MW; 9F309B5AE4E2AC1E CRC64;

Query Match 57.7%; Score 511; DB 1; Length 382;
Best Local Similarity 58.3%; Pred. No. 1.9e-36;
Matches 102; Conservative 30; Mismatches 43; Indels 0; Gaps 0;

OY 4 LBSGEVIVANTDQALGRSLAPRKITLTKGIGAGSKPELGRSAEQKVDIORM 63
Db 35 LQGVDFIVNTDQALHLSKAEVRLQGVKLTTRIGAGANPEIGKKAEESEHQIEALQ 94
OY 64 DSNMLFITGGMGGTCTGAPYVASARELIGLVGVSTPFRSGPRTLANAGV 123

```


[illegible]

```

OY      170  IIRPGLINL 178
DB      216  MVMGLINL 224

RESULT 5.
FTSZ_BORBU
ID      FTSZ_BORBU      STANDARD:      PRT:      399 AA.
AC      P45483; OS9183;
DT      01-NOV-1995 (Rel. 32, Created)
DT      15-DEC-1998 (Rel. 37, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Cell division protein ftsz.
GN      FTSZ OR BB0239.
OS      Borrelia burgdorferi ( Lyme disease spirochete).
OC      Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX      NCBI_faxID=139;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-ATCC 35210 / B31;
RX      MEDLINE=98065943; PubMed=9403685;
RA      Friser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA      Ladhari R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA      Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA      Peterson J., Keilavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA      van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA      Utenback T., Matthey L., McDonald L., Artlach P., Bowman C.,
RA      Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA      Smith H.O., Venter J.C.;
RT      "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT      burgdorferi."
RL      Nature 390:580-586(1997).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN-212;
RA      Ge Y., Old I.G., Saint-Girons I., Charon N.W.;
RL      Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN      [4]
RP      SEQUENCE OF 29-115 FROM N.A.
RC      STRAIN-212;
RX      MEDLINE=93146383; PubMed=1490605;
RA      Old I.G., MacDougall J.H., Saint-Girons I., Davidson B.E.;
RT      "Mapping of genes on the linear chromosome of the bacterium Borrelia
RL      burgdorferi: possible locations for its origin of replication."
RN      FEMS Microbiol. Lett. 78:245-250(1992).
RN      [5]
RP      SEQUENCE OF 1-115 FROM N.A.
RC      STRAIN-212;
RL      Old I.G.;
CC      Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: This protein is essential to the cell-division process.
CC      It seems to assemble into a dynamic ring on the inner surface of
CC      the cytoplasmic membrane at a place where division will occur,
CC      and the formation of the ring is the signal for septation to
CC      begin. Binds to and hydrolyzes GTP (By similarity).
CC      -1- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC      of the cytoplasmic membrane (By similarity).
CC      -1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: U43739; AAA85622.1; ALT_INIT.
DR EMBL: AE001137; AAC66649.1; ALT_INIT.
DR EMBL: X9685; CAA65464.1; ALT_INIT.
DR EMBL: L76303; AAB51402.1; -.
DR EMBL: Z12164; CAA78156.1; ALT_INIT.
DR HSSP: 057816; 1FSZ.
DR TIGR: B80299; -.
DR InterPro: IPR000158; FtsZ.
DR InterPro: IPR003008; Tubulin_FtsZ.
DR Pfam: PF00091; tubulin_1.
DR PRINTS: PR00423; CELDIVISFtsZ.
DR TIGRPFAMs: TIGR00065; ftsz_1.
DR PROSITE: PS01134; FtsZ_1; 1.
DR PROSITE: PS01135; FtsZ_2; 1.
DR Cell division; Septation; GTP-binding; Complete proteome.
KW NP_BIND 113 121 GTP (POTENTIAL).
FT CONFLICT 115 115 G -> A (IN REF. 4 AND 5).
FT CONFLICT 249 250 AA -> RR (IN REF. 3).
FT CONFLICT 336 336 T -> A (IN REF. 3).
SQ SEQUENCE 399 AA; 42398 MW; E808E36343E583 CRC64;

```

```

Query Match 57.2%; Score 506; DB 1; Length 399;
Best Local Similarity 58.7%; Pred. No. 5.3e-36;
Matches 101; Conservative 27; Mismatches 44; Indels 0; Gaps 0;

```

```

QY 7 VEFIVANTDQALGRSLAPKHTLTKDITKGLGAGSKPELGRSAEQKVDIORMLODSN 66
DB 47 VEFIVANTDQALGRSLAPKHTLTKDITKGLGAGSKPELGRSAEQKVDIORMLODSN 106
QY 67 MEFTFGMGSGGTGGAAPVAVASVARELGITVGVSPTPRSEGPRTLRANAGYKELAKY 126
DB 107 MEFTFGMGSGGTGGAAPVAVASVARELGITVGVSPTPRSEGPRTLRANAGYKELAKY 166
QY 127 VDTLIVPNQMLALADKSTMLFAFRYADVLLGKGVDTLIVPGLINL 178
DB 167 VDTLIVPNQMLALADKSTMLFAFRYADVLLGKGVDTLIVPGLINL 218

```

```

RESULT 6
FTSZ_CAUOCR STANDARD; PRT; 508 AA.
AC P52976;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cell division protein ftsz.
GN FTSZ OR CC2540.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CB15N / NA100;
RX MEDLINE=968270535; PubMed=8692812;
RA Ouardokus E., Din N., Brun Y.V.;
RT "Cell cycle regulation and cell type-specific localization of the
RL ftsz division initiation protein in Caulobacter."
RN Proc. Natl. Acad. Sci. U.S.A. 93:6314-6319(1996).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Niemman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heldelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeJoy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathevan J., Etmolova M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RN Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).

```

```

CC -1 FUNCTION: This protein is essential to the cell-division process.
CC It seems to assemble into a dynamic ring on the inner surface of
CC the cytoplasmic membrane at the place where division will occur,
CC and the formation of the ring is the signal for septation to
CC begin. Binds to and hydrolyzes GTP (by similarity).
CC -1 SUBUNIT: Aggregates to form a ring-like structure (by similarity).
CC -1 SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC of the cytoplasmic membrane (by similarity).
CC -1 SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL: U40273; AAC44223.1; -.
DR EMBL: AE005922; AAC24511.1; -.
DR HSSP: Q57816; 1FSZ.
DR TIGR: CC2540; -.
DR InterPro: IPR000158; FtsZ.
DR InterPro: IPR003008; Tubulin_FtsZ.
DR Pfam: PF00091; tubulin_1.
DR PRINTS: PR00423; CELDIVISFtsZ.
DR TIGRPFAMs: TIGR00065; ftsz_1.
DR PROSITE: PS01134; FtsZ_1; 1.
DR PROSITE: PS01135; FtsZ_2; 1.
KW Cell division; Septation; GTP-binding; Complete proteome.
FT NP_BIND 107 115 GTP (POTENTIAL).
FT CONFLICT 236 236 G -> A (IN REF. 1).
SQ SEQUENCE 508 AA; 54207 MW; 4E817C09P58C4CE CRC64;

```

```

Query Match 57.1%; Score 505; DB 1; Length 508;
Best Local Similarity 58.3%; Pred. No. 8.5e-36;
Matches 102; Conservative 28; Mismatches 45; Indels 0; Gaps 0;

```

```

QY 4 LEGVEIVANTDQALGRSLAPKHTLTKDITKGLGAGSKPELGRSAEQKVDIORMLO 63
DB 38 LEGVEIVANTDQALGRSLAPKHTLTKDITKGLGAGSKPELGRSAEQKVDIORMLO 97
QY 64 DSNMFLITGGMGGGTGGAAPVAVASVARELGITVGVSPTPRSEGPRTLRANAGYKEL 123
DB 98 GAHWFTTAGGGGTGGAAPVAVASVARELGITVGVSPTPRSEGPRTLRANAGYKEL 157
QY 124 AKYVDLIVPNQMLALADKSTMLFAFRYADVLLGKGVDTLIVPGLINL 178
DB 158 QRYVDLIVPNQMLALADKSTMLFAFRYADVLLGKGVDTLIVPGLINL 212

```

```

RESULT 7
FTSZ_BACSU STANDARD; PRT; 382 AA.
AC P17865;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cell division protein ftsz.
GN FTSZ.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=89008108; PubMed=3139638;
RA Beall B., Lowe M., Lutkenhaus J.;
RT "Cloning and characterization of Bacillus subtilis homologs of
RL Escherichia coli cell division genes ftsz and ftsA."
RN J. Bacteriol. 170:4855-4864(1988).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=168;

```


[illegible]

```

Db      36 KGVGEFIITANDVQALRKSKAEFTVIQLGPKYTRIGLGAASQPEYVGKAAEESQAIRELD    95
QY      64 DSNMLFTTGKGGGCTCGAAPVASVARRELGLTVGVSTPFRSEGNRRTRLANAGVEL    123
       : ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      96 GADMIFITAGMGGGTGTGAAPIVAGIAKEALGALTGVYTREPTEFGPKRGFFAAGIARL    155
QY      124 AKYVDILIVPNQNLLALADKSTIMLEAFRADVDLVLGVGVDYLIPRGLINTL    178
       : ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      156 KENDDTLLISNNRLLEVDDKKTPMLEAFREADNVLRGGVQGISDILTAPGIYNL    210

RESULT 10
FITSZ_ENTFE
ID     FITSZ_ENTFE   STANDARD:      PRT:      412 AA.
AC     008439;
DT     15-JUL-1998 (Rel. 36, Created)
DT     15-JUL-1998 (Rel. 36, Last sequence update)
DE     15-JUN-2002 (Rel. 41, Last annotation update)
Pr     Cell division protein fitsz.
CN     FITSZ.
OS     Enterococcus faecalis (Streptococcus faecalis).
OC     Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX     NCBI_TaxID=1351;
RN     [1]
RC     SEQUENCE FROM N.A.
RA     Pucci M.J., Thanaasi J.A., Discotto L.F., Kessler R.E.,
RL     Dougherty T.J.;
Submitted (MAR-1997) to the EMBL/Genbank/DDBJ databases.
CC     -! FUNCTION: This protein is essential to the cell-division process.
CC     It seems to assemble into a dynamic ring on the inner surface of
CC     the cytoplasmic membrane at the place where division will occur,
CC     and the formation of the ring is the signal for separation to
CC     begin. Binds to and hydrolyzes GTP (By similarity).
CC     -! SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC     -! SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC     of the cytoplasmic membrane (By similarity).
CC     -! SIMILARITY: BELONGS TO THE FITSZ FAMILY.
-----
Cc      This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc      between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc      the European Bioinformatics Institute. There are no restrictions on its
Cc      use by non-profit institutions as long as its content is in no way
Cc      modified and this statement is not removed. Usage by and for commercial
Cc      entities requires a license agreement (see http://www.isb-sib.ch/announce/
Cc      or send an email to license@isb-sib.ch).
-----
DR      EMBL: U94707; AAC45639.1; -.
DR      HSSP: O57816; IFSZ.
DR      InterPro: IPR000158; FtsZ.
DR      InterPro: IPR003008; Tubulin_FtsZ.
DR      Pfam: PF00091; tubulin_1.
DR      PRINTS: PRO0423; CEBLDVISFTSZ.
DR      TRIPFAMS: TIGR00065; ftsz_1.
DR      PROSITE: PS01134; FITSZ_1; FALSE_NEG.
DR      PROSITE: PS01135; FITSZ_2; 1.
DR      Cell division; Separation; GTP-binding.
KW      NP_BIND          105      113      GTP (POTENTIAL).
FT
SQ      SEQUENCE        412 AA; 44355 MW; 42EA85BAA70EF51F CRC64;

Query Match              55.3%; Score 489; DB 1; Length 412;
Best Local Similarity    56.6%; Pred. No. 1.6e-34;
Matches 99; Conservative 28; Mismatches 48; Indels 0; Gaps 0;

QY      4 LEVEEFIVANDPCALGRSLAPHKITTTGDKDTTGKLGAGSKPELGKRSAGQKDIORMQ    63
       : ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      36 VKGEVFITANTDVQALKHKAETIYIQLPKTRNLWGSGSPDEVQKAAEESQYIISLSQ    95
QY      64 DSNMLFTTGKGGGCTCGAAPVASVARRELGLTVGVSTPFRSEGNRRTRLANAGVEL    123
       : ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      96 GADMIFITAGMGGGTGTGAAPVAVAKIAKEALGALTGVYTREPTEFGPKRGFFAEGIALL    155
QY      124 AKYVDILIVPNQNLLALADKSTIMLEAFRADVDLVLGVGVDYLIPRGLINTL    178
       : ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      156 KENDDTLLISNNRLLEVDDKKTPMLEAFREADNVLRGGVQGISDILTAPGIYNL    210

```

```

DB      156 KENVYDILLISNNRLLLEVVDKKTPEMLLEAFREADNVLROGVGCSIDLITAPGYNL 210
      : |||::: | | : || | |||| |::: ||::: || | |
RESULT 11
FTSZ_MYCKA STANDARD; PRT; 386 AA.
ID      FTSZ_MYCKA
AC      Q9KHA25;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Cell division protein ftsZ.
GN      FTSZ.
OS      Bacterium kanasii.
OC      Bacteria: Actinobacteria: Actinobacteria (class): Actinobacteridae:
OC      Actinomycetales: Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_Taxid=1768;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Ramanujam S., Ajikumar P.;
RT      ftsZ gene of M. kanasii.;
RL      Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: This protein is essential to the cell-division process.
CC      It seems to assemble into a dynamic ring on the inner surface of
CC      the cytoplasmic membrane at the place where division will occur,
CC      and the formation of the ring is the signal for septation to
CC      begin. Binds to and hydrolyzes GTP (By similarity).
CC      -1- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC      of the cytoplasmic membrane (By similarity).
CC      -1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
EMBL: AF273451; AAF8784.2; -
DR      HSSP: Q57816; ftsZ.
DR      InterPro: IPR000158; ftsZ.
DR      InterPro: IPR003008; Tubulin_FtsZ.
DR      Pfam: PF00091; tubulin; 1.
DR      PRINTS: PR00423; CELLDIVISFSZ.
DR      TIGRFAWS: TIGR00065; ftsZ; 1.
DR      PROSITE: PS01134; FTSZ_1; 1.
DR      PROSITE: PS01135; FTSZ_2; 1.
KW      Cell division; Septation; GTP-binding.
FT      NP_BIND 101 109 GTP (POTENTIAL).
SQ      SEQUENCE 386 AA; 39051 MW; 215DE0B814ED593 CRC64;

Query Match 54.9%; Score 486; DB 1; Length 386;
Best Local Similarity 54.9%; Pred. No. 2.7e-34;
Matches 96; Conservative 33; Mismatches 46; Indels 0; Gaps 0;

OY      4 LEQVEFIANTDQALGRSLAPHKITLKGDTKGLGAGSKPELGRSAEQKVDIDRMQ 63
      : ||||| | | | | | : | : | : | : | : | : | : | : | : | : |
DB      32 LKGEVFIANTDQALMSDADVKLDVGRDSTRIGAGADPEVGRKAAADADIEELR 91
      : ||||| | | | | | : | : | : | : | : | : | : | : | : | : |
OY      64 DSNMLETGSGGCTGCAAPVAVASVARELGLTVGVVSTPERSGPNFTRLANAGVKEL 123
      : ||||| | | | | | : | : | : | : | : | : | : | : | : | : |
DB      92 GDMVFETAGEGGCTGTGAPVAVASVARELGLTVGVVSTPERSGPNFTRLANAGVKEL 151
      : ||||| | | | | | : | : | : | : | : | : | : | : | : | : |
OY      124 AKVVDLIIVPNQNLALADKSTMLFAFRADVLLEGVGVKVDLIIVPGLNL 178
      : ||||| | | | | | : | : | : | : | : | : | : | : | : | : |
DB      152 RESCDLIIVPNQNLALADKSTMLFAFRADVLLEGVGVKVDLIIVPGLNL 206
      : ||||| | | | | | : | : | : | : | : | : | : | : | : | : |

RESULT 12
FTSZ_STRCO STANDARD; PRT; 399 AA.
ID      FTSZ_STRCO

```

```

AC      P45500;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Cell division protein ftsZ.
GN      FTSZ OR SCO2082 OR SC4A10.15c.
OS      Streptomyces coelicolor.
OC      Bacteria: Actinobacteria: Actinobacteria (class): Actinobacteridae:
OC      Actinomycetales: Streptomycetaceae; Streptomycetaceae; Streptomyces.
OX      NCBI_Taxid=1902;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      STRAIN-A3(2);
RX      MEDLINE=95131746; PubMed=7830569;
RT      McCormick J.R., Su E.P., Driks R.;
RT      "Growth and viability of Streptomyces coelicolor mutant for the cell
RT      division gene ftsZ.";
RL      Mol. Microbiol. 14:243-254(1994);
RN      [2]
RP      SEQUENCE FROM N.A.
RA      STRAIN-A3(2) / M145;
RX      MEDLINE=21996410; PubMed=1200953;
RA      Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA      Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA      Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA      Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA      Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA      Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA      Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA      Warren T., Wietzorrek A., Woodward J., Barrill B.G., Parkhill J.,
RA      Hopwood D.A.;
RT      "Complete genome sequence of the model actinomycete Streptomyces
RT      coelicolor A3(2).";
RL      Nature 417:141-147(2002).
CC      -1- FUNCTION: This protein is essential to the cell-division process.
CC      It seems to assemble into a dynamic ring on the inner surface of
CC      the cytoplasmic membrane at the place where division will occur,
CC      and the formation of the ring is the signal for septation to
CC      begin. Binds to and hydrolyzes GTP (By similarity).
CC      -1- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC      of the cytoplasmic membrane (By similarity).
CC      -1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
EMBL: U10879; AAD10533.1; -
DR      EMBL: AL109663; CAB51991.1; -
DR      HSSP: Q57816; ftsZ.
DR      InterPro: IPR000158; ftsZ.
DR      InterPro: IPR003008; Tubulin_FtsZ.
DR      Pfam: PF00091; tubulin; 1.
DR      PRINTS: PR00423; CELLDIVISFSZ.
DR      TIGRFAWS: TIGR00065; ftsZ; 1.
DR      PROSITE: PS01134; FTSZ_1; 1.
DR      PROSITE: PS01135; FTSZ_2; 1.
KW      Cell division; Septation; GTP-binding; Complete proteome.
FT      NP_BIND 101 109 GTP (POTENTIAL).
SQ      SEQUENCE 399 AA; 41095 MW; EAD52B04CFEB4D39 CRC64;

Query Match 54.9%; Score 486; DB 1; Length 399;
Best Local Similarity 53.1%; Pred. No. 2.7e-34;
Matches 93; Conservative 38; Mismatches 44; Indels 0; Gaps 0;

OY      4 LEQVEFIANTDQALGRSLAPHKITLKGDTKGLGAGSKPELGRSAEQKVDIDRMQ 63
      : ||||| | | | | | : | : | : | : | : | : | : | : | : | : |
DB      32 LKGEVFIANTDQALMSDADVKLDVGRDSTRIGAGADPEVGRKAAADADIEELR 91
      : ||||| | | | | | : | : | : | : | : | : | : | : | : | : |

```

QY 64 DSNMFLITGGMGGCTGTGAPVAVASVARELGLTVGVSTPFRSEGNRTRLANAGKEL 123
 Db 92 GADWVFVYAGEGGCTGTGAPVAVANARSLGALTITGVTFPFEEGRRRANQADGJAEI 151
 QY 124 AKYVDLIVPNQNLALADKSTMTLEAFRYADVLLEGVGYTDLVRGLINL 178
 Db 152 REEDVTIVIPNDRLRLISIDROVSVLDARFSAQVLLSGVGLITDITTPGLINL 206

RESULT 13
 FTSZ_MYCTU STANDARD; PRT; 379 AA.
 ID FTSZ_MYCTU 008378;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cell division protein ftsZ.
 GN FTSZ OR RV2150C OR MT2209 OR MTCY270.18.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 MEDLINE=98295987; PubMed=634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holtroyd S.,
 Hornsby T., Jagels K., Krogh A., McLean A., Moule S., Murphy L.,
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.";
 Nature 393:537-544(1998).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White-O.,
 Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 Kolonay J.F., Nelson W.C., Umayam L.A., Ermolova M.D., Salzberg S.L.,
 Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,
 Bishai W.;
 "Whole genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains.";
 Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RL [1]
 CC -1- FUNCTION: This protein is essential to the cell-division process.
 CC It seems to assemble into a dynamic ring on the inner surface of
 CC the cytoplasmic membrane at the place where division will occur,
 CC and the formation of the ring is the signal for septation to
 CC begin. Binds to and hydrolyzes GTP (By similarity).
 CC -1- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
 CC of the cytoplasmic membrane (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: Z95388; CA086643.1;
 CC EMBL: AE007066; AAK46493.1; ALT_INIT.
 CC HSSP: Q57816; ftsZ.
 CC TIGR: MT2209;
 CC Tuberculist: RV2150C;
 CC InterPro: IPR000158; ftsZ.
 CC

DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; Tubulin_1.
 DR PRINTS: PR00423; CELLDVISTSZ.
 DR TIGRFAMs: TIGR0065; ftsZ_1.
 DR PROSITE: PS0134; FTSZ_1; 1.
 DR PROSITE: PS0135; FTSZ_2; 1.
 DR Cell division; Septation; GTP-binding; Complete proteome.
 FT NP_BIND 101 109 GTP (POTENTIAL).
 SQ SEQUENCE 379 AA; 38756 MW; 3F58035307878A9 CRC64;

Query Match 54.8%; Score 485; DB 1; Length 379;
 Best Local Similarity 54.3%; Pred. No. 3,1e-34;
 Matches 95; Conservative 35; Mismatches 45; Indels 0; Gaps 0;

QY 4 IEQVEIVANTDQALGRSLAPKRTKGLDGAGKPELGRSAEQGVQVQRMQ 63
 Db 32 LKGVFEIINTDAQALMSDADVKLVGRSTGLGAGAPVEYRKAAEDKDEIELR 91

QY 64 DSNMFLITGGMGGCTGTGAPVAVASVARELGLTVGVSTPFRSEGNRTRLANAGKEL 123
 Db 92 GADWVFVYAGEGGCTGTGAPVAVASTARKIGALTGVTFPFEEGRRRANQADGJAEI 151

QY 124 AKYVDLIVPNQNLALADKSTMTLEAFRYADVLLEGVGYTDLVRGLINL 178
 Db 152 RESCDTIVIPNDRLRLISIDROVSVLDARFSAQVLLSGVGLITDITTPGLINL 206

RESULT 14
 FTSZ_STRGR STANDARD; PRT; 407 AA.
 ID FTSZ_STRGR AC P45501;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cell division protein ftsZ.
 GN FTSZ.
 OS Streptomyces griseus.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 NCBI_TaxID=1911;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B2682;
 RX MEDLINE=94374704; PubMed=8086545;
 RA "Pharmatlake A.", Kendrick K.E.;
 RT "Expression of the division-controlling gene ftsZ during growth and
 RT sporulation of the filamentous bacterium Streptomyces griseus.";
 RL Gene 147:21-28(1994).
 CC -1- FUNCTION: This protein is essential to the cell-division process.
 CC It seems to assemble into a dynamic ring on the inner surface of
 CC the cytoplasmic membrane at the place where division will occur,
 CC and the formation of the ring is the signal for septation to
 CC begin. Binds to and hydrolyzes GTP (By similarity).
 CC -1- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
 CC of the cytoplasmic membrane (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: 007344; AAA56889.1;
 CC HSSP: Q57816; ftsZ.
 CC InterPro: IPR000158; FtsZ.
 CC InterPro: IPR003008; Tubulin_FtsZ.
 CC Pfam: PF00091; tubulin_1.
 CC PRINTS: PR00423; CELLDVISTSZ.
 CC TIGRFAMs: TIGR0065; ftsZ_1.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2003, 06:27:41 ; Search time 72 Seconds

509.394 Million cell updates/sec

Title: US-09-770-509-2

Sequence: 1 ASQLEGVETIVANTDQALG.....LLEGVKVTDLIVRPLINL 178

Scoring table: BLOSUM62

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summerless

Database

```

1: SP_RHEM1.21:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.podent:*
13: sp.virus:*
14: sp.vertebrate:*
15: sp.unclassified:*
16: sp.virus:*
17: sp.archaeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Length	DB	ID	Description
	1	547	61.8	253	2	08RMK5	08rmk5 azospirillum
	2	527	59.5	336	2	09RNN2	09rnn2 zymomonas m
	3	527	59.5	401	10	09M7M6	09m7m6 mallomonas m
	4	525	59.3	421	2	09AQ38	09aq38 ehrlichia c
	5	524	59.2	452	2	09AO37	09ao37 rickettsia
	6	521	58.9	400	2	09AC37	09ac37 anaplasma p
	7	520	58.8	357	16	08R9H2	08r9h2 thermomonas
	8	520	58.8	370	2	085474	085474 clostridium
	9	517	58.4	231	2	P77997	P77997 wolbachia s
	10	517	58.4	289	2	008390	008390 wolbachia s
	11	515	58.2	289	2	008392	008392 wolbachia s
	12	514	58.1	289	2	008471	008471 wolbachia s
	13	514	58.1	289	2	008389	008389 wolbachia s
	14	514	58.1	289	2	008391	008391 wolbachia s
	15	510	57.6	391	16	08Y5M5	08y5m5 listeria m
	16	510	57.6	392	16	0929Y5	0929y5 listeria m

17	507	57.3	373	16	097IE9	097IE9 clostridium
18	503	56.8	581	2	069074	069074 bartonella
19	499	56.4	590	2	069075	069075 bartonella
20	497	56.2	350	2	09F7C6	09F7C6 wolbachia e
21	496	56.0	581	2	09X6M9	09X6M9 bartonella
22	495	55.9	381	16	08XJUI	08XJUI clostridium
23	493	55.7	372	2	085475	085475 clostridium
24	493	55.7	417	16	09CEH2	09CEH2 lactococcus
25	491	55.5	351	2	09L7D8	09L7D8 wolbachia s
26	491	55.5	351	2	09L6V2	09L6V2 wolbachia e
27	490	55.4	407	10	09SSV6	09SSV6 cyanidiosch
28	489	55.3	392	2	09FIM6	09FIM6 shewanella
29	488	55.1	419	2	09ZAJ1	09ZAJ1 lactococcus
30	487	55.0	559	16	09AKB9	09AKB9 rhizobium l
31	486.5	55.0	318	2	086036	086036 wolbachia e
32	486	54.9	290	2	085473	085473 epulopiscu
33	485	54.9	402	2	085717	085717 streptomyce
34	485	54.8	348	2	086037	086037 wolbachia e
35	482	54.5	583	16	08UDN5	08UDN5 agrobacteri
36	481	54.4	373	2	09RME1	09RME1 rhizobium l
37	476	53.8	343	16	0986C2	0986C2 rhizobium l
38	476	53.8	338	16	08Y160	08Y160 bruceella me
39	475	53.7	566	2	0937A1	0937A1 bruceella ab
40	472	53.3	419	2	09ZHB9	09ZHB9 streptococ
41	471	53.3	419	16	097PE9	097PE9 streptococ
42	471	53.2	336	2	09Z3B8	09Z3B8 wolbachia s
43	469	53.0	331	2	09EVS3	09EVS3 wolbachia s
44	468	52.9	331	2	09EVS4	09EVS4 wolbachia s
45	468	52.9	334	2	09EVS0	09EVS0 wolbachia s

ALIGNMENTS

RESULT 1

ID	Q8RMK5	PRELIMINARY;	PRT;	253	AA
----	--------	--------------	------	-----	----

DT 01-JUN-2002 (Tremblere, 21, Created)

DT 01-JUN-2002 (Tremblay, 21, Last annotation update)

GN
GN
FTSZ.
GN

OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;

OX	NCBI_TaxID=192;
PN	[11]

RE SEQUENCE FROM N.A.
BC STBAIN=CD:

RT "A dd1B (D-alanine-D-alanine ligase) Tn5 mutant of Azospirillum

RT and maize root colonization.";

DR EMBL; AF492457; AAM1652.1; -.

SQ SEQUENCE 253 AA; 26638 MW; 13B8B38B0AF4702D CRC64;

Query Match	Score	DB 2;	Length
61.88;	547;	253	

Matches 107; Conservative 30; Mismatches 40; Indels 0; Gaps 0

2 S0LEGEVFIVANTDCQALGRSLAPHKRTTLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61

DB 36 SNLEGVDFVVGNIIDAGALKGSLCEKRVQLGTIMIRGLGAGSKPDVGKASAEQLLEELIGH 95

0Z LQDSNMLE I I GMSGG I CIGAF V VASVAREGDI I VGV VSI F F NSEGFNK I KLMNAG V K 121

22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051

— — — — —

DB 156 ELQOYDTLLITIPNQNFRANEXTTAFADAFRAADVLHSGVGYTDLNMPGLINL 212

RESULT 2

09RNN2 PRELIMINARY; PRT: 336 AA.

AC 09RNN2: 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE Cell division protein ftsz.

GN FTSZ.

OS Zymomonas mobilis.

OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;

OC Zymomonas.

OX NCBI_TaxID=542;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN-2M4:

RA Um H.W., Kang H.S.;

RT "Sequence analysis of 41E10 fosmid clone of Zymomonas mobilis.";

RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.

CC ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR, AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).

CC -1- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).

CC OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.

DR EMBL: AF179611; AD53930.1; -.

DR HSSP: Q57816; 1FSZ.

DR InterPro: IPR000158; FtsZ.

DR InterPro: IPR003008; Tubulin_FtsZ.

DR Pfam: PF00091; tubulin; 1.

DR PRINTS: PR00423; CELLDVIFTSZ.

DR TIGRFAMs: TIGR00065; ftsz; 1.

DR PROSITE: PS01134; FtsZ_1; 1.

DR PROSITE: PS01135; FtsZ_2; 1.

KW Cell division; GTP-binding; Septation.

KW SEQUENCE 336 AA; 35725 MW; 8CEPFLF7150436C7 CRC64;

SO Query Match

Best Local Similarity 59.5%; Score 527; DB 2; Length 336;

Matches 106; Conservative 29; Mismatches 43; Indels 0; Gaps 0;

QY 1 ASQEGVEFVANTDQALGSLAPHKITLTKGLGAGSKPELGKRSAGQKVDIOR 60

DB 38 ASGVGVFVIVNTDQALNISPAGRIQLGPTTGGAGSRREVGAAMAEETIQIOE 97

QY 61 MLDQSNMFLITGGMGCGTCTGAPVVAVARLGLITVGVSTPPRSRGNRLANAGV 120

DB 98 ALEGANMCTIAGMGCGTCTGAPVVAVARLGLITVGVSTPPRSRGNRLANAGV 157

QY 121 KELAYVDTLIVPNQNLALADKSTMLAEFRYADVLEGVGVDTLIVRGLINL 178

DB 158 EELQKHVDLITVIPNQNFRANEXTTAFADAFRAADVLHSGVGYTDLNMPGLINL 215

RESULT 3

09M7M6 PRELIMINARY; PRT: 401 AA.

AC 09M7M6: 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE Cell division protein ftsz.

GN FTSZ OR MSFTSZ-MT.

OS Mallomonas splendens.

OC Eukaryota; stramenopiles; Chrysophyceae; Synurales; Mallomonas.

OX NCBI_TaxID=52352;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN-MUCC 294;

RX MEDLINE=20144160; PubMed=10678836;

RA Beech P.L., Neu T., Schultz T., Herbert S., Lithgow T., Gilson P.R., McFadden G.I.;

RT "Mitochondrial ftsz in a Chromophyte Alga.";

RL Science 287:1276-1279(2000).

CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.

CC ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR, AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).

CC -1- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).

CC OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.

DR EMBL: AF120116; AAF35432.1; -.

DR HSSP: Q57816; 1FSZ.

DR InterPro: IPR000158; FtsZ.

DR InterPro: IPR003008; Tubulin_FtsZ.

DR Pfam: PF00091; tubulin; 1.

DR PRINTS: PR00423; CELLDVIFTSZ.

DR TIGRFAMs: TIGR00065; ftsz; 1.

DR PROSITE: PS01135; FtsZ_2; 1.

KW Cell division; GTP-binding; Septation.

KW SEQUENCE 401 AA; 42337 MW; E7B6A08C34A754BB CRC64;

SO Query Match

Best Local Similarity 59.5%; Score 527; DB 10; Length 401;

Matches 104; Conservative 27; Mismatches 47; Indels 0; Gaps 0;

QY 1 ASQEGVEFVANTDQALGSLAPHKITLTKGLGAGSKPELGKRSAGQKVDIOR 60

DB 100 AKRLSGVEFVCANTDQHLSTCLTENKTLQLGKSTQGLCGANPESGRRARAESEKRIAR 159

QY 61 MLDQSNMFLITGGMGCGTCTGAPVVAVARLGLITVGVSTPPRSRGNRLANAGV 120

DB 160 YIADANMFTIAGMGCGTCTGAPVVAVARLGLITVGVSTPPRSRGNRLANAGV 219

QY 121 KELAYVDTLIVPNQNLALADKSTMLAEFRYADVLEGVGVDTLIVRGLINL 178

DB 220 RSLQKVDLITIPNQNFRANEXTTAFADAFRAADVLHSGVGYTDLNMPGLINL 277

RESULT 4

09A038 PRELIMINARY; PRT: 421 AA.

AC 09A038: 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE Cell division protein ftsz.

GN FTSZ.

OS Ehrlichia chaffeensis.

OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;

OX Anaplasmataceae; Ehrlichia.

OX NCBI_TaxID=945;

RN (1)

RP SEQUENCE FROM N.A.

RA Lee K.N., Massung R.F., Padmalayam I., Baumstark B.;

RT "Characterization of the ftsz gene in Ehrlichia chaffeensis, the HGE agent and Rickettsia rickettsii.";

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.

CC ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR, AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).

CC -1- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).

CC OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.

DR EMBL: AF221944; AAK00615.2; -.

DR HSP: 057816; ftsz.
 DR Interpro: IPR000158; ftsz.
 DR Pfam: PF00091; tubulin_fts.
 DR PRINTS: PRO0423; CELDIVISFTSZ.
 DR TIGRPFAM: TIGR00065; ftsz. 1.
 DR PROSITE: PS01134; ftsz. 1; 1.
 DR PROSITE: PS01135; ftsz. 2; 1.
 DR Cell division: GTP-binding; Septation.
 DR SEQUENCE 421 AA; 45664 MW; E73CE4B1536255FA CRC64;

Query Match 59.3%; Score 525; DB 2; Length 421;
 Best Local Similarity 60.5%; Pred. No. 5.3e-39;
 Matches 107; Conservative 26; Mismatches 44; Indels 0; Gaps 0;

QY 2 SOLEGEVFIYVANTDQALGRSLAPHKITTKGDKITGAGSKPELGRSAEQKVDIORM 61
 DB 36 SLHGVNFVYVANDQALDCLSLSEKTIQGLGAGSLPEVGRGAEESEIETAE 95
 QY 62 LQDSNMLFTTGAGGGGTCTGAAPVAVASARELGLITGVVSTPFRSGPNTRLNAGVK 121
 DB 96 ISDSNMLFTTGAGGGGTCTGAAPVAVASARELGLITGVVSTPFRSGPNTRLNAGVK 155
 QY 122 ELAKYVDLIVPNQNLALADKSTMLAEFRYADVLEGVKVTDLVIRPGLINTL 178
 DB 156 ELQRYVDLIVIPNQLFRANENTTFADAKFLADTVLHGTGVRGTDLMVMPGLINTL 212

RESULT 5
 Q9AQ36 PRELIMINARY; PRT; 452 AA.
 ID 09AQ36:
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Cell division protein ftsz.
 GN ftsz.
 OS Rickettsia rickettsii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsiidae; Rickettsia.
 ON NCBI_TaxID=783;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lee K.N., Messing R.F., Padmalayam I., Baumstark B.;
 RT "Characterization of the ftsz gene in Ehrlichia chaffeensis, the HGE
 RT agent and Rickettsia rickettsii."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
 CC ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
 CC THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
 CC AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
 CC BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
 CC -1- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
 CC OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
 DR HSP: AF221945; AAK00617.1; -.
 DR EMBL: AF221945; 1FSZ.
 DR Interpro: IPR000158; ftsz.
 DR Interpro: IPR003008; tubulin_fts.
 DR Pfam: PF00091; tubulin_1.
 DR PRINTS: PRO0423; CELDIVISFTSZ.
 DR TIGRPFAM: TIGR00065; ftsz. 1.
 DR PROSITE: PS01134; ftsz. 1; 1.
 DR PROSITE: PS01135; ftsz. 2; 1.
 DR Cell division: GTP-binding; Septation.
 DR SEQUENCE 452 AA; 48365 MW; SDC6A1568630EBE0 CRC64;

Query Match 59.2%; Score 524; DB 2; Length 452;
 Best Local Similarity 56.2%; Pred. No. 7.1e-39;
 Matches 100; Conservative 38; Mismatches 40; Indels 0; Gaps 0;
 QY 1 ASOLEGEVFIYVANTDQALGRSLAPHKITTKGDKITGAGSKPELGRSAEQKVDIORM 60

DB 35 SANLOGANFVYVANTDQALDCLSLSEKTIQGLGAGSLPEVGRGAEESEIETAE 94
 QY 61 MLODSNMLFTTGAGGGGTCTGAAPVAVASARELGLITGVVSTPFRSGPNTRLNAGVK 120
 DB 95 YLENSNMVFTTGAGGGGTCTGAAPVAVASARELGLITGVVSTPFRSGPNTRLNAGVK 154
 QY 121 ELAKYVDLIVPNQNLALADKSTMLAEFRYADVLEGVKVTDLVIRPGLINTL 178
 DB 155 ELQRYVDLIVIPNQLFRANENTTFADAKFLADTVLHGTGVRGTDLMVMPGLINTL 212

RESULT 6
 Q9AQ37 PRELIMINARY; PRT; 400 AA.
 ID 09AQ37:
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Cell division protein ftsz.
 GN ftsz.
 OS Anaplasma phagocytophilum (Ehrlichia phagocytophila).
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Anaplasmataceae; Anaplasma.
 ON NCBI_TaxID=948;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lee K.N., Messing R.F., Padmalayam I., Baumstark B.;
 RT "Characterization of the ftsz gene in Ehrlichia chaffeensis, the HGE
 RT agent and Rickettsia rickettsii."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
 CC ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
 CC THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
 CC AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
 CC BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
 CC -1- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
 CC OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
 DR HSP: AF221945; AAK00616.1; -.
 DR EMBL: AF221945; 1FSZ.
 DR Interpro: IPR000158; ftsz.
 DR Interpro: IPR003008; tubulin_fts.
 DR Pfam: PF00091; tubulin_1.
 DR PRINTS: PRO0423; CELDIVISFTSZ.
 DR TIGRPFAM: TIGR00065; ftsz. 1.
 DR PROSITE: PS01134; ftsz. 1; 1.
 DR PROSITE: PS01135; ftsz. 2; 1.
 DR Cell division: GTP-binding; Septation.
 DR SEQUENCE 400 AA; 42001 MW; CDE8730E876FDC9B CRC64;

Query Match 58.9%; Score 521; DB 2; Length 400;
 Best Local Similarity 59.3%; Pred. No. 1.1e-38;
 Matches 105; Conservative 30; Mismatches 42; Indels 0; Gaps 0;

QY 2 SOLEGEVFIYVANTDQALGRSLAPHKITTKGDKITGAGSKPELGRSAEQKVDIORM 61
 DB 36 SLHGVNFVYVANDQALDCLSLSEKTIQGLGAGSLPEVGRGAEESEIETAE 95
 QY 62 LQDSNMLFTTGAGGGGTCTGAAPVAVASARELGLITGVVSTPFRSGPNTRLNAGVK 121
 DB 96 ISDSNMLFTTGAGGGGTCTGAAPVAVASARELGLITGVVSTPFRSGPNTRLNAGVK 155
 QY 122 ELAKYVDLIVPNQNLALADKSTMLAEFRYADVLEGVKVTDLVIRPGLINTL 178
 DB 156 ELQRYVDLIVIPNQLFRANENTTFADAKFLADTVLHGTGVRGTDLMVMPGLINTL 212

RESULT 7
 Q8R9H2 PRELIMINARY; PRT; 357 AA.
 AC Q8R9H2;

DR InterPro; IPR003008; Tubulin_FtsZ

Matches	100;	conservative	31;	Mismatches	40;	Indels	12;	Gaps	1;
---------	------	--------------	-----	------------	-----	--------	-----	------	----

QY 2 S0LBEEF1YANTDQALGRSLAPHKLTLLGKDITKGLGAGSKPELGRKRSAEQOKVDIORM 61
| : | | : | | | | | : | | | : | | | | | | : | | | : | | : |

AC 008471;
 DT 01-JUL-1997 (TREMblrel. 04, Created)
 DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Cell division protein ftsz (Fragment).
 GN FTSZ.
 OS Wolbachia sp.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Wolbachiae; Wolbachia.
 NCBI_TaxID=956;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M36, AND 1148;
 RA Jager C.R., Pintureau B., Heddi A.;
 RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
 CC ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
 CC THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
 CC AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
 CC BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
 CC -1- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
 CC OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
 CC EMBL: U95754; AAB54072.1; -;
 CC EMBL: U95750; AAB54068.1; -;
 CC HSSP: O57816; 1FSZ.
 DR InterPro: IPR000158; FtsZ.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin; 1.
 DR PRINTS: PR00423; CELLDVIFTSZ.
 DR TIGRFS: TIGR00065; ftsz.1.
 DR PROSITE: PS01134; FTSZ_1; 1.
 DR PROSITE: PS01135; FTSZ_2; 1.
 KW Cell division; GTP-binding; Septation.
 FT NON_TER 1
 FT NON_TER 289
 SQ SEQUENCE 289 AA; 30701 MW; BF0D6BEBCB430A07 CRC64;
 Query Match 58.1%; Score 514; DB 2; Length 289;
 Best Local Similarity 56.1%; Pred. No. 3.2e-38;
 Matches 106; Conservative 29; Mismatches 42; Indels 12; Gaps 1;
 QY 2 SOLEGEFFVANTDCOALGRSLAPHKITITGKDTKGLGAGSKPELGRSAEQKVDIQM 61
 DB 2 SNLQGVNFVANTDQALGKSLCKDKITQGLNITKGLGAGALPDIGKAEEISIDETMEH 61
 QY 62 LQDSNMLFITGGMGGTCTGAAPVAVASVAREL-----GILTVGVSTPFRSEG 109
 DB 62 IRDSHMLFITGGMGGTCTGAAPVAVASVAREL-----GILTVGVSTPFRSEG 109
 QY 110 PNRTRLANAGVRELKAYVDLIVPNONLALADKSTTMEAFRYADVLLGKGVYDL 169
 DB 122 VRRMRTAELGLELQKYVDLIVPNONLFRANKEKTFADAFOLADNVHLIGIRGYDL 181
 QY 170 IVRPGILNL 178
 DB 182 MIMPGILNL 190
 RESULT 13
 008389 PRELIMINARY; PRT; 289 AA.
 AC 008389;
 DT 01-JUL-1997 (TREMblrel. 04, Created)
 DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Cell division protein ftsz (Fragment).
 GN FTSZ.
 OS Wolbachia sp. 1032.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Wolbachiae; Wolbachia.
 NCBI_TaxID=59578;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1032;
 RA Jager C.R., Pintureau B., Heddi A.;
 RT "Comparison between phylogenetic trees of some Trichogramma species
 and their Wolbachia endosymbionts."
 RL Russ. Entomol. J. 7:163-168(1998).
 CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
 CC ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
 CC THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
 CC AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
 CC BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
 CC -1- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
 CC OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
 CC EMBL: U95749; AAB54067.1; -;
 CC HSSP: O57816; 1FSZ.
 DR InterPro: IPR000158; FtsZ.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin; 1.
 DR PRINTS: PR00423; CELLDVIFTSZ.
 DR TIGRFS: TIGR00065; ftsz.1.
 DR PROSITE: PS01134; FTSZ_1; 1.
 DR PROSITE: PS01135; FTSZ_2; 1.
 KW Cell division; GTP-binding; Septation.
 FT NON_TER 1
 FT NON_TER 289
 SQ SEQUENCE 289 AA; 30702 MW; B50A70FF77E41111 CRC64;
 Query Match 58.1%; Score 514; DB 2; Length 289;
 Best Local Similarity 56.1%; Pred. No. 3.2e-38;
 Matches 106; Conservative 29; Mismatches 42; Indels 12; Gaps 1;
 QY 2 SOLEGEFFVANTDCOALGRSLAPHKITITGKDTKGLGAGSKPELGRSAEQKVDIQM 61
 DB 2 SNLQGVNFVANTDQALGKSLCKDKITQGLNITKGLGAGALPDIGKAEEISIDETMEH 61
 QY 62 LQDSNMLFITGGMGGTCTGAAPVAVASVAREL-----GILTVGVSTPFRSEG 109
 DB 62 IRDSHMLFITGGMGGTCTGAAPVAVASVAREL-----GILTVGVSTPFRSEG 109
 QY 110 PNRTRLANAGVRELKAYVDLIVPNONLALADKSTTMEAFRYADVLLGKGVYDL 169
 DB 122 VRRMRTAELGLELQKYVDLIVPNONLFRANKEKTFADAFOLADNVHLIGIRGYDL 181
 QY 170 IVRPGILNL 178
 DB 182 MIMPGILNL 190
 RESULT 14
 008391 PRELIMINARY; PRT; 289 AA.
 AC 008391;
 DT 01-JUL-1997 (TREMblrel. 04, Created)
 DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Cell division protein ftsz (Fragment).
 GN FTSZ.
 OS Wolbachia sp. MB35.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Wolbachiae; Wolbachia.
 NCBI_TaxID=59581;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB35;
 RA Jager C.R., Pintureau B., Heddi A.;
 RT "Comparison between phylogenetic trees of some Trichogramma species
 and their Wolbachia endosymbionts."
 RL Russ. Entomol. J. 7:163-168(1998).
 CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
 CC ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF

CC THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
 CC AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
 CC BEGIN. BINDS TO AND HYDROLYSES GTP (BY SIMILARITY).
 CC -1 SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
 CC -1 SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
 CC OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
 CC -1 SIMILARITY: BELONGS TO THE FTSZ FAMILY.
 DR EMBL: U95752; AAB54070.1; -.
 DR HSP: Q57816; FTSZ.
 DR InterPro: IPR000158; Ftsz.
 DR InterPro: IPR003008; Tubulin_Ftsz.
 DR Pfam: PF000691; tubulin_1.
 DR PRINTS: PR00423; CELDVISFTSZ.
 DR TIGRFAMs: TIGR00065; ftsz_1.
 DR PROSITE: PS01134; FTSZ_1; 1.
 DR PROSITE: PS01135; FTSZ_2; 1.
 KM Cell division, GTP-binding; Septation.
 FT NON_TER 1 1
 FT NON_TER 289 289
 SQ SEQUENCE 289 AA; 30700 MW; 3D874344EB4DE409 CRC64;
 Query Match 58.1%; Score 514; DB 2; Length 289;
 Best Local Similarity 56.1%; Pred. No. 3.2e-38;
 Matches 106; Conservative 29; Mismatches 42; Indels 12; Gaps 1;
 QY 2 SOLEGEFFIANTDCOALGRSLAPRHKITLGDITKLGAGSKPELGRSAEQKVDIORM 61
 DB 2 SNLGVFVAVNTDQALAEKSLCKKIQGLNLTGKLGAGLPDYGKAAESIDIEIMH 61
 QY 62 LQDSNMLFTGGMGGCTGAPVAVASVAREL-----GLTVGVSTPRSEG 109
 DB 62 IRDSHMLFTAGMGCGTGTGAAPVIAKAAEAVAVKDKAKEKKITLVGVVTKPFEG 121
 QY 110 PNRTPLANAGVKELAKYVDTLIVPNQNLALADKSTMLFAFRYADVDLLEGVKGVTDL 169
 DB 122 VRMRRTAELEERLQKTYVDTLIVPNQNLFRANKEKTFADAFQLADNVHLHIGRGVTDL 181
 QY 170 IVRRGLINL 178
 DB 182 MIMPGLINL 190
 RESULT 15
 Q8Y5M5 PRELIMINARY; PRT; 391 AA.
 AC Q8Y5M5;
 DT 01-MAR-2002 (Tremblrel. 20. Created)
 DT 01-MAR-2002 (Tremblrel. 20. Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21. Last annotation update)
 DE ftsz protein.
 GN FTSZ OR LMO2032.
 OS Listeria monocytogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BGD-E / SEROVAR 1/2A;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Baquero F., Frangeul L., Buchrieser C., Rusnlok C., Amend A.,
 RA Baquero F., Berche P., Blocher H., Brandt P., Chakraborty T.,
 RA Chardit A., Cherouni F., Couve E., de Daruvar A., Denoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Entian K.-D., Falhl W., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordloek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
 RT "Comparative genomes of Listeria species."
 RL Science 294:849-852(2001).
 DR EMBL: AL591982; CAD00110.1; -.

DR ListList; LMO2032; -.
 DR InterPro: IPR000158; Ftsz.
 DR InterPro: IPR003008; Tubulin_Ftsz.
 DR Pfam: PF000691; tubulin_1.
 DR PRINTS: PR00423; CELDVISFTSZ.
 DR TIGRFAMs: TIGR00065; ftsz_1.
 DR PROSITE: PS01134; FTSZ_1; 1.
 DR PROSITE: PS01135; FTSZ_2; 1.
 KM Complete proteome.
 SQ SEQUENCE 391 AA; 41350 MW; 6B9F75B2B1629C58 CRC64;
 Query Match 57.6%; Score 510; DB 16; Length 391;
 Best Local Similarity 56.0%; Pred. No. 1.1e-37;
 Matches 98; Conservative 35; Mismatches 42; Indels 0; Gaps 0;
 QY 4 LEGVEFIYANTDCOALGRSLAPRHKITLGDITKLGAGSKPELGRSAEQKVDIORM 63
 DB 35 VQGVFISVNTDQALAEKSLCKKIQGLNLTGKLGAGLPDYGKAAESIDIEIMH 94
 QY 64 DSNMLFTGGMGGCTGAPVAVASVARELGLTVGVSTPRSEGPNRTPLANAGVKEL 123
 DB 95 GSDMVFTAGMGCGTGTGAAPVIAQIAKEMGALTGVVTRPFEGEKRRTQALTGTEAM 154
 QY 124 AKYVDTLIVPNQNLALADKSTMLFAFRYADVDLLEGVKGVTDLIVRRGLINL 178
 DB 155 KEAVDTLIVIPNDRLLOIVDKNTPMLEAFREADNVLRQGVQGISDLIAVPELLINL 209
 Search completed: June 2, 2003, 08:41:12
 Job time : 85 secs

DR PROSITE: PS01134; FTSZ_1; 1.
 DR PROSITE: PS01135; FTSZ_2; 1.
 KW Cell division; Septation; GTP-binding.
 FT NP_BIND 101 109 GTP (POTENTIAL).
 SQ SEQUENCE 407 AA; 41809 MW; 01D2898E01322005 CRC64;

Query Match 54.8%; Score 485; DB 1; Length 407;
 Best Local Similarity 53.1%; Pred. No. 3,4e-34;
 Matches 93; Conservative 38; Mismatches 44; Indels 0; Gaps 0;

QY 4 LGEVEFIVANTDQALGRSLAPHKITLTKGLGAGSKPELGKRSADQKVDIQRLQ 63
 Db 32 LKVEEFTALNTDQALMTKADRVIGLVNTEGSLGAGSQPEVGRMAAEBCIDEIDHLN 97
 QY 64 DSNMLFITGSGGGTCTGAAPVVASVARELGILTVGVSTPFSEGNPRTRLANAGVEL 123
 Db 92 GDMVFTAGGGGTCTGAAPVVASVARELGILTVGVSTPFSEGNPRTRLANAGVEL 151
 QY 124 AKYVDLIVPNONLLALADKSTTMEAFRYADVDLLEGVKGVTDLIVPGLINL 178
 Db 152 REEVDTLIVIPNDRLISDRQSVLDFAKSDQVLLSGVGITDLYTTPGLINL 206

RESULT 15

FTSZ_AGRU STANDARD; PRT; 583 AA.

AC 030992;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cell division protein ftsz.

GN FTSZ.

OS Agrobacterium tumefaciens.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Rhizobium.

OX NCBI_TaxID=358;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-A136;

RX MEDLINE=98012980; PubMed=9352931;

RA Ma X., Sun O., Wang R., Singh G., Jonietz E.L., Margolin W.;

RT "Interactions between heterologous FtsA and FtsZ proteins at the FtsZ ring."

RL J. Bacteriol. 179:6788-6797(1997).

-1- FUNCTION: This protein is essential to the cell-division process. It seems to assemble into a dynamic ring on the inner surface of the cytoplasmic membrane at the place where division will occur, and the formation of the ring is the signal for septation to begin. Binds to and hydrolyzes GTP (By similarity).

-1- SUBUNIT: Aggregates to form a ring-like structure (By similarity).

-1- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface of the cytoplasmic membrane (By similarity).

-1- SIMILARITY: BELONGS TO THE FTSZ FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----

CC EMBL: AF024659; AAC45821.1; -

CC HSSP: 057816; 1FSZ.

CC InterPro: IPR000158; FtsZ.

CC InterPro: IPR003008; Tubulin_FtsZ.

CC Pfam: PF00091; tubulin.1.

CC PRINTS: PR00423; CELDIVISFTSZ.

CC TIGRFAms: TIGR00065; ftsz.1.

DR PROSITE: PS01134; FTSZ_1; 1.

DR PROSITE: PS01135; FTSZ_2; 1.

DR PROSITE: PS01135; FTSZ_2; 1.

DR PROSITE: PS01135; FTSZ_2; 1.

DR PROSITE: PS01135; FTSZ_2; 1.

SQ SEQUENCE 583 AA; 61936 MW; 3F415B5FBCA1CB CRC64;

Query Match 54.5%; Score 482; DB 1; Length 583;

Best Local Similarity 57.1%; Pred. No. 9e-34;

Matches 100; Conservative 26; Mismatches 49; Indels 0; Gaps 0;

QY 4 LGEVEFIVANTDQALGRSLAPHKITLTKGLGAGSKPELGKRSADQKVDIQRLQ 63
 Db 38 LGVDEIVANTDQALMTKADRVIGLVNTEGSLGAGSQPEVGRMAAEBCIDEIDHLN 97
 QY 64 DSNMLFITGSGGGTCTGAAPVVASVARELGILTVGVSTPFSEGNPRTRLANAGVEL 123
 Db 98 GTHMCFVTAGGGGTCTGAAPVVASVARELGILTVGVSTPFSEGNPRTRLANAGVEL 157
 QY 124 AKYVDLIVPNONLLALADKSTTMEAFRYADVDLLEGVKGVTDLIVPGLINL 178
 Db 158 QKSVDTLIVIPNDRLISDRQSVLDFAKSDQVLLSGVGITDLYTTPGLINL 212

Search completed: June 2, 2003, 08:39:32
 Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 2, 2003, 08:47:19 ; Search time 2357 Seconds

(without alignments)
2197.836 Million cell updates/sec

Title: US-09-770-509-2

Perfect score: 885
Sequence: 1 ASQLEGEVFIANTDQALG.....LLEGVKGVTDLIVRPLINL 178

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+g2n.model -DEV-xlh
-Q=/cgn2_1/USPTO.spool/US09770509/runat_27052003_083648_18820/app_query.fasta_1.337
-DB-GenEmbl -OPMT-fastap -SUFFIX-rge -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS-bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.cdl -THR=5
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE-LOCAL
-OUTFMT-pio -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US09770509_@CGN_1_1_2466_@runat_27052003_083648_18820 -NCPU=6 -ICPU=3
-NO_XLPRY -NO_MMAP -LNSEQUENT -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=130 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_hum:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	547	61.8	3824	1	AF492457
2	527	59.5	1404	8	AF120116
3	527	59.5	38510	1	AF179611
4	525	59.3	1542	1	AF221944
5	524	59.2	1540	1	AF221946
6	524	59.2	8328	1	AE008653
7	521	58.9	1474	1	AF221945
8	520	58.8	1202	1	AF067822
9	520	58.8	11683	1	AE013119
10	518	58.5	279110	1	RFXR03
11	517	58.4	637	1	WSU37260
12	517	58.4	868	1	WSU95751
13	515	58.2	868	1	WSU95753
14	514	58.1	868	1	WSU95749
15	514	58.1	868	1	WSU95750
16	514	58.1	868	1	WSU95752
17	514	58.1	868	1	WSU95754
18	511	57.7	303249	1	AE001515
19	510	57.6	213251	6	AX413015
20	510	57.6	258650	1	AL596171
21	510	57.6	295050	1	AL591982
22	510	57.6	349980	6	AX417047
23	509	57.5	950	6	AX436249
24	509	57.3	13165	1	AE007679
25	507	57.2	8834	1	BBCDG
26	506	57.2	15272	1	BORFSA
27	506	57.2	34817	1	BORFSA
28	506	57.2	44380	1	AE001137
29	506	57.1	1877	1	CCU40273
30	505	57.1	10487	1	AE005922
31	503	56.8	1864	1	AF061746
32	503	56.8	1864	6	AX191760
33	503	56.7	1134	6	AX431765
34	502	56.7	1893	1	AF061747
35	498	56.4	4022	1	BACFSA
36	498	56.3	208230	1	BSUB0008
37	497	56.2	1052	1	AF282845
38	496	56.0	1859	1	AF141018
39	495	55.9	296750	1	AP003191
40	495	55.8	1782	1	AF007266
41	494	55.7	1201	1	AF067823
42	493	55.7	13423	1	AE006416
43	492	55.6	11551	1	EHY13922
44	492	55.5	1055	1	AF220605
45	491	55.5	1055	1	AF220605

RESULT 1

ALIGNMENTS

LOCUS	AF492457	3824 bp	DNA	linear	BCT 11-APR-2002
DEFINITION	Azospirillum brasilense D-alanine:D-alanine ligase (ddlB), partial cds, and cell division fts gene cluster, partial sequence.				
ACCESSION	AF492457				
VERSION	AF492457.1				
KEYWORDS	GI:20136382				
SOURCE	.				
ORGANISM	Azospirillum brasilense. Azospirillum brasilense. Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae; Azospirillum.				
REFERENCE	1 (bases 1 to 3824)				
AUTHORS	Jofre,E.C., Lagares,A. and Mori,G.B.				
TITLE	A ddlB (D-alanine-D-alanine ligase) Tn5 mutant of Azospirillum brasilense is altered in exopolysaccharide production, salt tolerance, and maize root colonization				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 3824)				
AUTHORS	Jofre,E.C., Lagares,A. and Mori,G.B.				
TITLE	Direct Submission				
JOURNAL	Submitted (14-MAR-2002) Natural Sciences, Universidad Nacional Rio Cuarto, Route 36 KM 601, Rio Cuarto, Cordoba 5800, Argentina				
FEATURES	Location/Qualifiers				
source	1..3824				
gene	/organism="Azospirillum brasilense"				
	/strain="Cd"				
	/db_xref="taxon:192"				
CDS	<1..632				
	/gene="ddlB"				
	<1..632				
gene	/gene="ddlB"				
	/codon_start=-3				
	/transl_table=11				
CDS	/product="D-alanine:D-alanine ligase"				
	/protein_id="AAM11649.1"				
	/db_xref="GI:20136383"				
gene	/translation="KALAPGVSPKGLVITKGLTGCAGHPAPRIYVPGDSITYGTVLRGQNSPYGDAMTFEERALVEEFTIGRELTVGWSGLDECPVTEIFRPAQVVDYPAKYSAGAAVHTVPAQIDPEHVAEAKRLAVLHRLGCGVSRSDFRMDRSKGTGDFPLEINQPGMTPLSLVEQAAVSGLTGALVAWVWENNAICOLD"				
	611..1579				
	/gene="ftsQ"				
CDS	617..1579				
	/gene="ftsQ"				
	/note="may be involved in septum formation"				
gene	/codon_start=-1				
	/transl_table=11				
	/product="cell division protein FtsQ"				
CDS	/protein_id="AAM11650.1"				
	/db_xref="GI:20136384"				
	/translation="MSRRLSDPQFRFGSGSANRGDDPVPYPRAMAKSAQCKNRRAMRWTRPAYKAAIILITPVLYVAGMAAMQRTIATTAALQSLTQTSASQAFATA DVLVEGETDPASILTLVIGVQSGDPLTAVTLDACKELSLPVASXSIERHLPDILFVLEERQEPMAIWOHDKRFTVIDREGRPADATELARGNRRIETLPQVVGANPMQVVKPLKLAIDNVPALREKVSASVSGDRMDKLNGVYVKKLEAPAMGALROLAEMDATGVLLADIVVAIDLRONDRAVLQTSATVLPWTEENKKRPGKKT"				
gene	1662..2924				
	/gene="ftsA"				
	1662..2924				
CDS	/gene="ftsA"				
	/note="may be involved in anomalous filament growth and a component of the septum; may interact with FtsZ"				
	/codon_start=-1				
gene	/transl_table=11				
	/product="cell division protein FtsA"				
	/protein_id="AAM11651.1"				
CDS	/db_xref="GI:20136385"				
	/translation="MGFNCAKKPKRATRGIIAALDVGSTVCCILARVDDAGVRIV GICHQIIVAGAGAIIDMEAEFTSIGAAV				

gene	CDS	BASE COUNT	ORIGIN
gene	<p>VLTGASOLPCMRRLAOLILDKOVRIGPFTITGLENAGGSPSYSTAAGLLHVRN</p> <p>ADLPVAGHAGAGGTFGRVGLMRLN"</p> <p>3067. .>3824</p> <p>/gene="fts2"</p> <p>3067. .>3824</p> <p>/note="essential for the cell-division process; assembles into a dynamic ring on inner surface of cytoplasmic membrane at location of cell division; binds to and hydrolyzes GTP"</p> <p>/transl_start=1</p> <p>/codon_start=1</p> <p>/product="cell division protein ftsZ"</p> <p>/protein_id="AAM11652.1"</p> <p>/db_xref="GI:20136386"</p> <p>/translation="MINITIPSTPELKPRTITVGVGAGNAYNNIKTSMLSEVDYVGNIDAOALKSLCEKRVOLSTYMTRGIGASKRDVGASAEDELFTIGHLEBANWVFTMGGGTGTGAAPVATARAERGLITVGVNKPPEHFEAHMRGLNLSGIALDOYVDITLIIIPNMLFRIANKEKTFADFAKADVDLVHGVGVYDLWVMGLINLDRADIRSVMTENKRAMGTGEAGEERRAIGRRNPHLPAA"</p>	<p>664 a 1353 c 1260 g 544 t 3 others</p>	
Alignment Scores:			
Pred. No.:	9.42e-41	Length:	3824
Score:	547.00	Matches:	107
Percent Similarity:	77.40%	Conservative:	30
Best Local Similarity:	60.45%	Mismatches:	40
Query Match:	61.81%	Indels:	0
DB:	1	Gaps:	0
US-09-770-509-2 (1-178) x AF492457 (1-3824)			
OY	2	SerGlnLeuGlnGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg	21
DB	3172	TCCAACCTCGAAGCGGTGACTTGCTGTCGGCAACACGACGGCAGCCCTCAAGGC	3231
OY	22	SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly	41
DB	3232	TGCGTCTGCGAAGAGCGGCTCACCTCGGACACCATGAGCGCGGCTGGCGCGGC	3291
OY	42	SerLysProGluLeuGlyLysArgSerIleGluGlnGlnLysValAspIleGlnArgMet	61
DB	3292	TCCAAGCGCGAGTGGCGCGGCGCTCGCGCGAGAGGAGCTCGAGAGATATCGGTAC	3351
OY	62	LeuGlnAspSerAsnMetLeuPheIleThrGlyLysMetGlyGlyThrCysThrGly	81
DB	3352	CTCGAAGCGCGCAACATGATGTTCATCACCGCGCATGGCGGCGGACGCGACGGGT	3411
OY	82	AlaAlaProValAlaAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValAla	101
DB	3412	GGGCGCGCGGTATCGCCCGCGCGCGCGGCGGAGCGGCGGCTGTCGACCGTGGCGGTG	3471
OY	102	SerThrProPheArgSerGlnGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys	121
DB	3472	GCAAGCGCTTCCACTTCGAGGGCGGACCGCATGGCGGTGGCGCATCGCATGGCC	3531
OY	122	GluLeuAlaLysIleThrValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeu	141
DB	3532	GAGCTGACGAGTATGTTGACACCTGTCTATCCCGAACCGAACCCTGTCGCATC	3591
OY	142	AlaAspLysSerThrThrMetLeuGlnAlaPheArgThrArgAlaAspAspValLeuLeu	161
DB	3592	GCCACACGAGAACGACCTTGGCGGACCGCTTCAAGATGGCCGACGACGCTTCTCATTC	3651
OY	162	GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu	178
DB	3652	GGCGTGGCGGCGTGCAGCATGATGATGATGCCGCGCATCAACCTG	3702

mitochondrial protein, complete cds.

ACCESSION AF120116
VERSION AF120116.1 GI:7024509
KEYWORDS
SOURCE Mallomonas splendens.
ORGANISM Mallomonas splendens
REFERENCE Eukaryote; stramenopiles; Chrysophyceae; Synurales; Mallomonas.
AUTHORS Beech, P.L., Nheu, T., Schultz, T., Herbert, S., Litgow, T.,
Gillson, P.R. and McFadden, G.I.
TITLE Mitochondrial FtsZ in a chromophyte alga
JOURNAL Science 287 (5456), 1276-1279 (2000)
MEDLINE 20144160
PUBMED 10678836
REFERENCE 2 (bases 1 to 1404)
AUTHORS Beech, P.L., Schultz, T., Vuong, T., Gillson, P.R. and McFadden, G.I.
TITLE Direct Submission
JOURNAL Submitted (13-JUN-1999) School of Botany, University of Melbourne,
Gretton St, Parkville, Vic. 3052, Australia
FEATURES
source
1. 1404
/organism="Mallomonas splendens"
/strain="WUCC 294"
/db_xref="taxon:52552"
1. 1404
/gene="Msfcsz-mt"
10. 1215
/note="putative mitochondrial division protein"
/codon_start=1
/product="FtsZ"
/protein_id="AAF35432.1"
/translation="MRTGANKRIISLSPRIHSDGASLNAKPIRSVKPKPKQYDSR
SGNSQASTREHVAKDEKPEKRLRRTGETTEFLPKIVGVGGGCAVNMILAK
LSGVFCANTDAHLSTCLTENKQLGKESQGLCCGANPESGRRAAEKSEILARY
IADANWFIATMGSGGTGAAPVAECMEMDILVAAYTRFSEFGKRAKLANGL
IRSLDEVDLIIIPNONIFKLIANSTADADGLADILLGKSIITDCAVPLNG
LDFADVTVMSGMHAIIMGTOAEGEDRIIRANADLNPLDGGDSVSRAKGMVNI
TGKDLTLVEYDAAARITSEDEDANVIFSSPDESIGSIRSVATGILAPGA
AATAAPVIR"
BASE COUNT 404 a 278 c 373 g 349 t
ORIGIN
Alignment Scores:
Pred. No.: 2.05e-39 Length: 1404
Score: 527.00 Matches: 104
Percent Similarity: 73.608 Conservative: 27
Best Local Similarity: 58.438 Mismatches: 47
Query Match: 59.558 Indels: 0
Gaps: 0
US-09-770-509-2 (1-178) x AF120116 (1-1404)
QY 1 A l s e r g i n l e u g l y a l g l u p h e i l e v a l a l a s n t h r a s p c y s g l a l a l e u g l y 20
D b 307 G C T A G A A A C G T C T G A G A T T T G C T C G C G A A T A C G A C C A C A C A C A T C T T C A 366
QY 21 A r g s e r i u a l p r o h i l e y s i l e t h r l e u g l y l y s a s p i l e t h r l y s e g l y l e u g l y a l a 40
D b 367 A C T T G C T T G A C T G A A A C G C T T C A G T A G A A A A G A C C A C C A G A G A T T G G T T G C 426
QY 41 G l y s e r l y s p r o g l u l e u g l y l y s a r g s e r a l a g l u g l i n g l i l y s a l a s p i l e g l n a r g 60
D b 427 G G T G C A A A C C A G A T C A G G G C G T C G A C C C G G A A A G A A A G A A A G A A T T G C C A G A 486
QY 61 M e t l e u g l i n a s p e r a s e m l e u p h e i l e t h r g l y g l y m e t c l y g l y l y t h r c y s t h r 80
D b 487 T A T A T T C C A G A T C T A A T A T G T A T T A A C T G C C G G A T G G T G T G G A C A G A C A G A C A 546
QY 81 G l y a l a l a p r o v a l i v a l a s e r v a l a a r g l u l e u g l y l l e u t h r v a l g l y a l 100
D b 547 G G G C C G C G C T G T A G A C C G A G G T T T G C A T G A A A A G A C A T T C T A A C G G T G G C A G T G 606

QY 101 V a l s e r t h r p r o p h e a r g s e r g l u g l y p r o a s n a r g t h r a r g l e u a l a s n a l a c l y a l 120
D b 607 G T C A C T A A A C A T A T T G C T T T G A G G A G A C A T C G C C T G A G C A A A G A A G A A T A 666
QY 121 L y s g l u l e u a l a s t y r a l a s p t h r l e u i l e v a l a l p r o a n g l a s n l e u a l a 140
D b 667 A G C T C T C G A A G A T C G T G T T G A C C G C T A A T A T T A T T C C A A A T C A A A T A T A T T C A A G 726
QY 141 L e u a l a s p l y s e r t h r m e t l e u g l u a l a p h e a r g t y r a l a s a s p a l l e u l e u 160
D b 727 C T C A T T A A C G C G C G A C T C A T G C C G A T G C G T T G C C G C T G A C A G A C A T T T G T G T G 786
QY 161 G l u g l y a l l e u g l y a l t h r a s p l e u i l e v a l a r g p r o g l y l e u i l e a s l e u 178
D b 787 G C C G C G T A G A G A C A C G A C G A C T G A T T G C G C G G G A C A T G A C T T G 840

RESULT 3
AF179611/c
LOCUS
DEFINITION Zymomonas mobilis ZM4 fosmid clone 41E10, complete sequence.
ACCESSION AF179611
VERSION AF179611.1 GI:5834350
KEYWORDS
SOURCE Zymomonas mobilis.
ORGANISM Zymomonas mobilis.
REFERENCE 1 (bases 1 to 38510)
AUTHORS Um, H.W. and Kang, H.S.
TITLE Sequence analysis of 41E10 fosmid clone of Zymomonas mobilis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 38510)
AUTHORS Um, H.W. and Kang, H.S.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-1999) Microbiology, Seoul National University,
San 56-1 Shillimdong Kwanaku, Seoul 151-742, Korea
FEATURES
source
1. 38510
/organism="Zymomonas mobilis"
/strain="ZM4"
/db_xref="taxon:542"
/db_xref="fosmid 41E10"
1. 465
/gene="PPIASE"
1. 465
/gene="PPIASE"
/codon_start=1
/transl_table=11
/product="peptidyl prolyl cis trans isomerase"
/protein_id="AAD53917.1"
/db_xref="GI:5834351"
/translation="DPKGDGSGSSLPNLPALFNDMPHVKEGVSMARAQEKDSANSQF
FLVYQPTSELDEYTPIGRVYSGMEYDAVEKGPAPHPSIIVPASISGSDIHPTTE
EYKAAYEKOEAENLAIQAKOEVEVRNRRMSFLPNHARLLPMQNRRLRQGH"
502. 1530
/gene="queA"
502. 1530
/gene="queA"
/note="queA"
/codon_start=1
/transl_table=11
/product="S-adenosylmethionine tRNA ribosyltransferase"
/protein_id="AAD53918.1"
/db_xref="GI:5834352"
/translation="MRVDLFDLPELPESIALRPAVPRESARLLQVTSQEMDYHVAL
PMLQEGDILLIFNDRIIPAYLGGKKNQARISANMIRREGIRQQAFAVRNARKLSDD
IYDFGQDVSARVHRNDGSSVLEFEGKNEVEILERAGNPLPPIYAKRPADQODR
ODYOTIFAEKEGVAAPTASLHFTBELLOLEKGVKHTVTLLHVGAGTFLPMVEDS
DDHVMHSEKQISAEFRANAIINERKKOGGRIYARTGTSILRLDSAADEGLYHPSPSDY
DITFTGRIYFRYIDGLMTNHLPKSTLFMLVSALMGREKMLAAYQHAIDTGYREFSYG
DSSLLPK"
1596. 2174
gene

```

CDS
    /gene="fpra"
    /note="PMP oxidase"
    /codon_start=1
    /transl_table=11
    /product="pyridoxamine 5'-phosphate oxidase"
    /protein_id="A053919.1"
    /db_xref="GI:5834353"
    /translation="MTODPHKMFEDMDDEAKKATDDPTAMALATASKOAPPCVMMTLKTHMDGEVFTNLSRKHELLENVATLITFWKTRROYRIGATLTLSDEADA YRTRARKSOLGEMASBQSRPLPARDFEKRADIENAYESGDVPRPYWGFGRVSPIRNEFMDFRELRHREFTLINDRWOSEFLYP"
    complement(2214..3332)
    /gene="ntrc"
    complement(2214..3332)
    /gene="ntrc"
    complement(2214..3332)
    /codon_start=1
    /transl_table=11
    /product="nitrogen assimilation regulatory protein"
    /protein_id="A053920.1"
    /db_xref="GI:5834354"
    /translation="MASISFRRASVVENTALIGESPAMREIRYLIOQVAPPSAVLITGPSGKREMASSIHASSPSSGRFIPVKGCAIPVLLSEELFEGHEGDTAGTGOTCGREFADGGTLELDEIDMPFDMQVALLVLEENIITRVGKOSIHVDRTIISATHRIDHMALEHKAFFREDLEFRLAVFPIHLPSLADRSDDVPLAEYFCDRLPENKRFCSFSDSHLYSOXWMPGNIRELRNLIERTSILYPOQYGEVOTOKLNLAVSSQPDRLALMNLSSPSEISSDRSLADLIEDDPDESENLALFADPEIPGDVAKTELVELECRYITAAANSANNIVDAAARLLGLHRTTLEKNNKYSIS"
    4117..5646
    /gene="zmlorf1"
    4117..5646
    /gene="zmlorf1"
    /codon_start=1
    /transl_table=11
    /product="phosphatidylinositol glycan homolog"
    /protein_id="A053921.1"
    /db_xref="GI:5834355"
    /translation="MANKDKSAASFSLKSIOASHVTLVLLTLALCLMTAIFPIGYHDEIFQYLEOAHRLAFYGVYIPEWERYGMRGAVPULLAAMKMGDMNTOSFLYVLPDAITLALSGIVGSAVKGRLHSFPTGLAMVATWEIVLFAAHPISLSLFAAFMPAAVYLIAEEKKSLITSGFLMGLSALLRPYRPSILFTPLIRINROMKXFTLGGIAGVFSSISDIPIGEYPLMLFRNIDONLIFHAEDYGTAPITSYFFDIWTNKMWSLIPITLVAAIRPALFETALVNLIFHMLVSHKRYRTFSSGIFILAAIGSALLIEKATINFTPLHKLQAIAATAMGACSLSTGNISHWYQYRGVLAQAEKARSLNLCGLTLOIDWSESGYSLHRIPIYSLPSMYSRTYRDPITALSPQLSASNAIISAASSGKNLPKEKRYOVCYQKNGEKEDICLFRPSSCFKEKADNFEINKVLKNDIO"
    complement(5933..7819)
    /gene="zmlorf2"
    complement(5933..7819)
    /gene="zmlorf2"
    /codon_start=1
    /transl_table=11
    /product="porin homolog"
    /protein_id="A053922.1"
    /db_xref="GI:5834356"
    /translation="MKNFKRIARVAOWVIALSIATTSAAWADPLGMWLAKVNNHHQKSSASQVEODRDTNTVNTSPVANOPOAQTSQVSPSATPYQSVNRESQASOIPADADTNDVIRMSDNDNDNDSAGRVADAPOMAEVSHOHSPAPFPASLTJVGSELPEPTDRLPVGSGVVEENKGLDSTYOSTGRNAPAMIGVGNKSNVLAIRPKDRMFTWGLIKPMLRKGQIDIGSMGCVETAFNARGDGRHGLRTAOWLVASPDMERLGVARGKTFNMTKRGGOIEQAINSTLTLEMEVYGRNGIMRLIEGYSQOIRFNIVKGRITENNREGMAATGCSFMSLIGCGNAGNIVGDYWMPPVSNMAGRYTVQYSKEVITIGAVFVNNKNIQGLVLENNEMDPHPPEAGLYIGEVKMAPKFGRENLESLFKFDGWTDDKSYNLRVODVAMOTGYSSGCPFVRGAVYTAEOOVLRPDSHROTIVAALRKADGKNTSTAVSAAILFGVSKERPDVGGAGCTWTLSNAFLKQALSAONRHQEYSELFTNIQLSPAVFLRPNIQIVHPGGITVANGPVDTPYNNKNITLLGLKSNV"
    8520..10292
    /gene="kefc"
    8520..10292
    /gene="kefc"

```

```

    /note="kefc; K+ efflux antiporter"
    /codon_start=1
    /transl_table=11
    /product="glutathione-regulated potassium efflux system protein"
    /protein_id="A053923.1"
    /db_xref="GI:5834357"
    /translation="MNSLSFSDSVILGAAGLIIPAFARFRISPVIGTLACAVLGPKALGSLVEYPMWKHTITSDPGIEPDAFGVILMFSLISIEFSKRLMROAVFGITGRKLGGGLIAGLTYALGENWAGIAGTGSORIVNRRGAASSGNGTCRKPAILLFEDLALVPIIFALVPYHMLKGIINGTDLIRITNGGARVIALIVLGRALPTLRQAAARTSAVFLSISLVILASVATSAAGLSPVGAITAGLLAETDYHAEVENVILPEORLAIGVFLVMSGLRLIMANMOALAGVAVLNTVLTSTLRISHVRLATAVESCLIMASPESETTVIATARAQOLIMPESTAFMEIVAGLITPHLAKISHEVSLKLTLEHMLTLPSENEPIEDIONHTIVAGIRGKVIADVLTSEKKYLAIDSDIDVILNARDGYHATPGDISNPETIOGLKNNIAVYALINDPSVMVNVHRLHRTYDPLFTIARASRRHAQVFRAGADIVAFTESSLRALPATLISLGSVSTGHAVAAIHDKRASBRENLKAAEEGNMOAIRHTPSIMPEA"
    complement(10368..13034)
    /gene="alas"
    complement(10368..13034)
    /gene="alas"
    complement(10368..13034)
    /note="alas"
    /codon_start=1
    /transl_table=11
    /product="alanine-tRNA synthetase"
    /protein_id="A053924.1"
    /db_xref="GI:5834358"
    /translation="MITNTEIRSRSRIRFRKPSDSTIAPVDPNDPTLMPVNAVYPRKPTFGLESRPYKTAASSOKCVYRAGSKHNDLVGTYARHHTPEMIGNSFGCYFKERYIELAMGLITKEMGLDPERCLTYHTDEARNTLMKRIKGLPDRILIKATSDNFWMDGTGPGCGPCEIFGDPEIPGASLAEGBDFEIMLVEMQYQVNAETRLNLPRESIDTGMGLERIASVLCGVHNDYDPTDEKALIAAGSELHTATDQGFKASHRYADHLRAEGLVADGVLPANEGGVYLRIMRRAMRAHLIGAKEPLRLVAPALISEGMAYPELVRKALIEETLLETFREKOTLANGKITKDETHQKSGDTLPGAVAFRLYDTYGPYPLTADALARNITVQAGTDAMAEBRKARANKSGEASPEIPIEDIDQLGTEFTGYTRKSGOITLILDKGRVETAGOGDITITIMQTPYGGSGGKGDTGLVITGNDLKMVTDTOKPLGRIHAIATIEKEIITGDIQLOVDINHRNRLRANNSATHLHAALRLDQGHVSQKGSNVSNERLRFESHQRLADQELAEAVNQILNNSVVTTRPTEAVAGAMALFGEKYGEMRVFMSCLNDQSESSVYELCGTAVSALQOIGLTHIVSEIVASGIRIEKVEGTBEALMVGRRRLRETASILKVADEEPTRTAIIIDERRKCSLADOKRALPANAGCGGNNAPDKIGAYHTIGOVIGLEPKALGLTIDENKKTIESGVATLQVLENSRAVATIGVSLKCKISAVDVRRAVELTGKGGGPGDAQGGGPNGENAQAQLEAVALLERA"
    complement(13374..14303)
    /gene="zmlorf3"
    complement(13374..14303)
    /gene="zmlorf3"
    complement(13374..14303)
    /codon_start=1
    /transl_table=11
    /product="unknown"
    /protein_id="A053925.1"
    /db_xref="GI:5834359"
    /translation="MDKDDQHDHDMNSHGRPOENNTYVNPFRGNASHGSSRDTSFYDHSADADGDGEDSLPRLPEVEDDDNKGCIPIYIYALALLTALLGTLGYMFRHPPAVSGNASVITAEPEYTKPREGEKMTIEGSESYTASBEGDLSIDTSVQPEQPTIAKQAEQITPPPPASAPVTKRPDVNEPEMSDWTGDRHSHTAHPAPRAPAVKHEKEVPAHEHPYKAAREKPPAVVAPPAKVEKPEAEHQAQVAPKPSREDSVAGVSSLHNDHNPDEAKKAPAGCCYVGCILM"
    complement(14307..16049)
    /gene="args"
    complement(14307..16049)
    /gene="args"
    complement(14307..16049)
    /note="args"
    /codon_start=1
    /transl_table=11
    /product="arginyl-tRNA synthetase"
    /protein_id="A053926.1"
    /db_xref="GI:5834360"
    /translation="MDYFNRPDLSFPCFALSSIIDTFPAEGLLPAGLDRSRAVEBPPRDPSSGDLAINAAWVILAKAATAPALAKLIVESLSLQAGVESVETAGPGFINRLDDCMRKELTHILEGODYGRSDIGONLRVAVEYVSANPTGPMHGHGRGAVGALASLIDFAGYKVTRETYVNDAGOOVOTLARSAIRYREALGEITITPEGLIPGYLVFVGRLL

```

Alignment Scores:

Pred. No.: 9,2e-38 Length: 38510
Score: 527.00 Matches: 106
Percent Similarity: 75.84% Conservative: 29
Best Local Similarity: 59.55% Mismatches: 43
Query Match: 59.55% Indels: 0
DB: 1 Caps: 0

US-09-770-509-2 (1-178) x AF179611 (1-38510)

QY 1 A1aserGluLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGly 20
DB 22221 GCATCGAGGTGCGAGGCTGTCGATTATGTAGACCAATACAGATGCCACGCTTAAT 22162
QY 21 ArgSerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGly 40
DB 22161 AACTTCGGGCAACAGCGGATTCACCTGGCCGACGACGACGAGGCTTGGGGCA 22102
QY 41 GlySerLysProGluLeuGlyLysArgSerAlaGluGlnLysValAspIleGlnArg 60
DB 22101 GGATCGAGACCGGAGAGTGGGAAAGCGCCGACAGAAACGATCGACGATTCAGAA 22042
QY 61 MetLeuGlnAspSerAsnMetLeuPheIleThrGlyMetGlyGlyThrCysThr 80
DB 22041 GCCTTGGAAGCGCAGAAATGCTTATCGCCGCTGGATGGGGCGGCGACGCGCAG 21982
QY 81 GlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyVal 100
DB 21981 GGTGCGAGCTCGGTCATCGCAAAAGTACGCCGCTGATGCTGATTCGACAGTGGTGT 21922
QY 101 ValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
DB 21921 GTGACCAACACCGTTCAACTTTGAAGGTAAAGCGGCGGCGTCAAGTGAAGGCGATT 21862
QY 121 LysGluLeuAlaLysIleValAspThrLeuIleValValProAsnGlnAsnLeuAla 140
DB 21861 GAAGACTCTCAAAAGCATGTCATACGCTATGTTTCACAAACGAAATTTGTTTG 21802
QY 141 LeuAlaAspLysSerThrThrMetLeuGluAlaPheArgThrAlaAspAspValLeu 160
DB 21801 ATTGCCAATCCGACACGACGACTTAAAGCAGCCCTTCACAGATGGCGGATGAGCTTTG 21742
QY 161 GluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
DB 21741 CAGGGGTTCGCGGTATACCGATTTGATGTTGCCCGGCTTATCAATCTT 21688
RESULT 4
AF221944 1542 bp DNA linear BCT 15-MAY-2001
LOCUS Ehrlichia chaffeensis cell division protein Fltz2 (fltz2) gene,
DEFINITION complete cds.
ACCESSION AF221944
VERSION AF221944
KEYWORDS
SOURCE
ORGANISM Ehrlichia chaffeensis.
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Anaplasmataceae; Ehrlichia.
REFERENCE
AUTHORS Lee, K.N., Massung, R.F., Padmalayam, I. and Baumstark, B.
TITLE 1 (bases 1 to 1542)
JOURNAL Ehrlichia chaffeensis.
Unpublished
2 (bases 1 to 1542)
COMMENT Submitted (07-JAN-2000) Centers for Disease Control and Prevention,
1600 Clifton Rd., Decatur, GA 30329, USA
FEATURES
Location/Qualifiers
1..1542
/organism="Ehrlichia chaffeensis"

gene
CDS

/db_xref="taxon:945"
254..1519
/gene="fltz2"
254..1519
/gene="fltz2"
/codon_start=1
/transl_table=11
/product="cell division protein Fltz2"
/protein_id="AAK0615.2"
/db_xref="GI:14043017"

BASE COUNT
ORIGIN

577 a 232 c 280 g 453 t

Alignment Scores:

Pred. No.: 3.49e-39 Length: 1542
Score: 525.00 Matches: 107
Percent Similarity: 75.14% Conservative: 26
Best Local Similarity: 60.45% Mismatches: 44
Query Match: 59.32% Indels: 0
DB: 1 Gaps: 0

US-09-770-509-2 (1-178) x AF221944 (1-1542)

QY 2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
DB 359 TCTATCTCAGCGGCTTAACTTTGATGATGCTTAACTGATGATGATGATGATGATGAT 418
QY 22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyVal 41
DB 419 TCCCTTGCAAAAAGAAATTCAGTAGATGATGATGATGATGATGATGATGATGATGAT 478
QY 42 SerLysProGluLeuGlyLysArgSerAlaGluGlnLysValAspIleGlnArgMet 61
DB 479 TCACCTCCAGAAAGTGGTAGAGTGGTGCAGAAAGCAATAAATGATTAATGAGAA 538
QY 62 LeuGlnAspSerAsnMetLeuPheIleThrGlyMetGlyGlyThrCysThrGly 81
DB 539 ATCTCAGATAGCAATATGCTGATCAATACGCTGGAATGGTGAGGAGACAGAAAGCC 598
QY 82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyVal 101
DB 599 GCTGCTCCCTGTAATTCAGAGTGGCCAGAGCAACAAATTTTAACCATAGAGAGTGT 658
QY 102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 121
DB 659 ACAAAACCTTTCACCTTCAGAGGCGACATCGAATGACGACACAGAAATTTGGTTGAA 718
QY 122 GluLeuAlaLysIleValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeu 141
DB 719 GAATTAACAAGGATGATGATCTCATAGATTAATCTTAATCAAAATCTATTCAGAGAT 778
QY 142 AlaAspLysSerThrThrMetLeuGluAlaPheArgThrAlaAspAspValLeuGlu 161
DB 779 GCTAATGAAAAAACAACATTTGCGAGATTCAACTTGACAGACACTGCTTACATACA 838
QY 162 GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
DB 839 GGTGTCGTCGATCACTAATGATGATGATGATGATGATGATGATGATGATGAT 889
RESULT 5
AF221946 1540 bp DNA linear BCT 15-MAY-2001
LOCUS Rickettsia rickettsii cell division protein Fltz2 (fltz2) gene,
DEFINITION complete cds.
ACCESSION AF221946

```

VERSION      AF221946.2  GI:14043019
KEYWORDS
SOURCE       Rickettsia rickettsii.
ORGANISM     Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
              Rickettsiaceae; Rickettsiinae; Rickettsia; spotted fever group.
REFERENCE    1 (bases 1 to 1540)
AUTHORS      Lee, K.N., Massung, R.F., Padmalayam, I. and Baumstark, B.
TITLE        Characterization of the ftsZ gene in Ehrlichia chaffeensis, the HGE
              agent and Rickettsia rickettsii
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 1540)
AUTHORS      Lee, K.N., Massung, R.F., Padmalayam, I. and Baumstark, B.
TITLE        Direct Submission
JOURNAL      Submitted (07-JAN-2000) Centers for Disease Control and Prevention,
              1600 Clifton Rd., Decatur, GA 30329, USA
              On May 15, 2001 this sequence version replaced gi:12655831.
COMMENT      Location/Qualifiers
FEATURES     source
              1..1540
              /organism="Rickettsia rickettsii"
              /db_xref="taxon:783"
              /gene="ftsZ"
              /gene="ftsZ"
              /codon_start=1
              /product="cell division protein ftsZ"
              /protein_id="AAK00617.1"
              /transl_table=11
              /db_xref="GI:12655832"
              /translation="MVLNFKAPENIVLKPTITVGVGAGSNAVNMNISANLQGANFV
              VANTAOSLEHSICTNKIOLGYSTTRGLAGASPEVALAAESESERANLENMNV
              FTAGAGGCGTGSAPVARIKELGLTVYVTPFEGHMKRDLKLELOOF
              VTLIVIPNOLFRIPANBOTPFADFMADVDVLAHGRGTIDIMKGLINDPDIK
              AVMSMGKAMGTGASGSDRAKKAESAISNPLDSSMCAAGVILNITGSDMTL
              FEVDNAKRIREVLDNAILIGSTFNPGLKIRSVVATGIDADVPYKPIAT
              TTNIPEETYNKAIAQPTQIEPDPNSYSTKINETIDSPINOMLIGNEKELGHANT
              FNKSDDSKPSFLGKIMGSLRASNNQTLERKNVVSITLDQNKESDIDIPAFLRKO
              RD"
BASE COUNT   537 a      267 c      305 g      431 t
ORIGIN
Alignment Scores:
Pred. No.:    4,31e-39      Length:      1540
Score:         524.00      Matches:      100
Percent Similarity: 77.53%   Conservative: 38
Best Local Similarity: 56.18%  Mismatches: 40
Query Match:   59.21%      Indels:        0
DB:            1          Gaps:          0

US-09-770-509-2 (1-178) x AF221946 (1-1540)
OY      1  A laser glu leu glg l y val g l u phe lle val a la s n p h r a s p c y s g l a l a l e u g l y 20
DB      141 AGTGTAAATCTCAAGAGTCTAATTTGTAGTAGCTAATCTGATGACAAATCCCTTGAA 200
OY      21  Arg ser leu ala p r o h i l y s i l l e r h r l e u g l y l a s p l l e r h r l y s g l y l e u g l y a l a 40
DB      201 CATCTTTATGACACAAACAAATACAACTCGGTCTTACGACTAGACGCTTGCGGCA 260
OY      41  Gly ser l y s p r o g l u l e u g l y l y s a r g s e r a l a g l u g l n g l y s v a l a s p l l e g l n a r g 60
DB      261 GGAGCTTCTCCGAGGCTTGACACTTGCCTCAGCAAGAACGAAACGAAATTCGTAA 320
OY      61  Met leu g l n a s p s e r a s m e t l e u p h e l l e t h r g l y m e t g l y g l y t h r c y s t h r 80
DB      321 TACCTAGAAATAGCAATATGCTATTATTCGCGACGATGAGCGGTGCTACAGATCC 380
OY      81  G l y a l a a l a p r o v a l a l a s e r v a l a a r g l u l e u g l y l e l e u t h r v a l g l y a l 100
DB      381 GGTTCTGACCGCTATATGACGCAATTCGTAAGAACTAGCAATCCCTTACGCTTGCGGTA 440
OY      101 Val s e r t h r i p r o p h e a r g s e r g l u g l y p r o a s n a r g t h r a r g l e u a l a s n a l a g l y a l 120

```

```

DB      441 GTAACTAAACCTTCCATTTTGAAGCGGCTACATCGTATGAAACACGCGCATTAAGACTT 500
OY      121 Lys Glu Leu Ala l a s y r v a l a s p t h r l e u l l e v a l p r o a s n g l n a s n l e u a l a 140
DB      501 ATTGAATATGACGATTCGTTGATCTTAATTTGAAATACCAACCAAAATATTCCTCGT 560
OY      141 Leu Ala s p l y s s e r t h r t h r m e t l e u l l u a l p h e a r g t y r a l a s p a s p a l l e u 160
DB      561 ATTGCTATGACAAACAGACGCTTTGCCGATGCTTCAAAATGCAATGCAATGATGATTA 620
OY      161 Glu Gl y a l l y s g l y a l t h r a s p l e u l l e v a l a r g p r o g l y l e u l l e a s n l e u 178
DB      621 GCAGCGCTTAGAGAGATGACGATTTAATGATTAATGACCGGAGCTATTAATCTA 674
RESULT 6
LOCUS      AE008653
DEFINITION Rickettsia conorii Malish 7, section 85 of 114 of the complete genome.
ACCESSION AE008653
VERSION   AE006914
KEYWORDS  Rickettsia conorii.
          Rickettsia conorii.
          Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
          Rickettsiaceae; Rickettsiinae; Rickettsia; spotted fever group.
ORGANISM  Rickettsia conorii.
REFERENCE 1 (sites)
AUTHORS   Ogata, H., Audic, S., Barbe, V., Artiguenave, F., Fournier, P. E.,
          Raoult, D., and Claverie, J. M.
          Selfish DNA in protein-coding genes of Rickettsia
          Science 290 (5490), 347-350 (2000)
REFERENCE 2 (sites)
AUTHORS   Ogata, H., Audic, S., and Claverie, J.-M.
          Selfish DNA and the origin of genes
          Science 291 (5502), 252-253 (2001)
REFERENCE 3 (bases 1 to 8328)
AUTHORS   Ogata, H., Audic, S., Renesto-Audiffren, P., Fournier, P.-E., Barbe, V.,
          Samson, D., Roux, V., Cossart, P., Weissenbach, J., Claverie, J.-M. and
          Raoult, D.
          Mechanisms of evolution in Rickettsia conorii and R. prowazekii
          Science 293 (5537), 2093-2098 (2001)
TITLE     JOURNAL
          MEDLINE
          PUBMED
REFERENCE 4 (bases 1 to 8328)
AUTHORS   Ogata, H., Audic, S., Renesto-Audiffren, P., Fournier, P.-E., Barbe, V.,
          Samson, D., Roux, V., Cossart, P., Weissenbach, J., Claverie, J.-M. and
          Raoult, D.
          Direct Submission
          Submitted (26-APR-2001) Unite des Rickettsies, Faculte de Medecine,
          CNRS UMR 6020, Universite de la Mediterranee, 27 boulevard Jean
          Moulin, Marseille Cedex 05 13385, France
          A public version of R. conorii genome database is accessible at
          http://igs-server.cnr-mrs.fr/. The database intends to provide
          updated data. Annotation of the genome is an ongoing task whose
          goal is to make the genome sequence more useful. Comments to the
          authors are appreciated.
FEATURES  source
          1..8328
          /organism="Rickettsia conorii"
          /strain="Malish 7"
          /db_xref="taxon:781"
          /gene="ftsZ"
          /gene="ftsZ"
          /note="RC1015"
          /gene="ftsZ"
          /gene="ftsZ"
          /codon_start=1
          /transl_table=11
          /product="cell division protein ftsZ"
          /protein_id="AAU03553.1"

```

```
/db_xref="GI:15620131"
/translation="MVLNKAPENIIVLPTITVEGVGAGNAVNMMISANIGANFV
VANTDAOLESILCTNRIQLGVSTTRIGLAGASPEVGAIAOSESSEINYLENSMV
ETIAGMGGTGTSAPVIARIAKEGILTVGVTPFPHREGHMKMTADKGLIELOOF
VDFLIYIPNONIFRIANEORTFADAFKADVDLHAGVGYDPLMIPGLINDRADIK
AVNSEMKAMGTGEASGEDRAIKAESAINPLDHSMGARGVLLITGSDMTL
TEVDMANRIREEVDNLDANIIFGSTNPELKGMRVAVGADADKVPYKPAIAE
TTNIVPEETRIKALAOPTOAIEMPDFNSYSTEINIEITDPSINQNEIGNEKEIGLHANT
FNKSEDDSPKPSFGLKIMGSLRASNNQTLERKNVYVSTLDQNKESDIDIDIAFLRKK
RD"
gene
1835..2473
/feature="RC1016"
1835..2473
/feature="RC1016"
/codon_start=1
/translation="unknown"
/product="unknown"
/protein_id="AAL03554.1"
/db_xref="GI:15620132"
/translation="MEDKPYSLDIDGATYYAALDPTELNQNMMINTREPSATSLDFHL
IETLPNFSODKOKKALTEFELKQNNPAKITYLEQOQNSOILSINQNNNNYILF
PKIINLEEOYIDILNLLGLTGNKCNFEPSNNGSEIYFRLPNNERLMAKYPFLANO
PLIKENKFEISIELSTRKINENGVDVDSGFSNQWYKCFIYNSRRDNGF"
2660..2866
/feature="RC1017"
2660..2866
/feature="RC1017"
/codon_start=1
/translation="similarity to ecotin precursor"
/protein_id="AAL03555.1"
/db_xref="GI:15620133"
/translation="MQATKNAMODCNRYWFGGKLETKTLEGWYNYIIDQVSDHPAS
TMACPRVKATIDVSVFLGDET"
3006..3031
/feature="REP01, repeated element"
complement(3084..3656)
/feature="RC1018"
complement(3084..3656)
/feature="RC1018"
/codon_start=1
/translation="unknown"
/product="unknown"
/protein_id="AAL03556.1"
/db_xref="GI:15620134"
/translation="MFIOTEDTPNPDAITFFPGOISSDQVFFSEIAVFKRSKIAE
SLFYNNKSVFPGSDFTYTKOAEKNQVYKPELMTYIMDFVSGFFPEENTKADN
VNHLNGLSIEIKQIVETITRVRPSVAQDGDIIYKGFENGVALALGACILCPSS
TITLNGIESMLHFVEVQEVAVEEGFE"
complement(3852..3880)
/feature="REP01, repeated element"
3920..3957
/feature="REP08, short repeated element"
complement(4004..5068)
/feature="ampg2"
/feature="RC1019"
complement(4004..5068)
/feature="ampg2"
/codon_start=1
/translation="ampg protein"
/protein_id="AAL03557.1"
/db_xref="GI:15620135"
/translation="MNFTHIFKFLMGLPILKVPFAPLSKRGYKCLITIALVSCIPRY
VLPNPNHFIPEFALCLVAVAFSSSYDMLQSSQMLITKKNGISRAACTTCRI
GIIASGALYSTISMODVYRAMALICISLITIIYFKKOKIIINDREHWA
FDFVKPKMLIYSEMLYRLQDNFSLVPMMEFLDGYTKKDLALGKAKGCMATY
LGGFGLCKRKEYSYLKRVLIIHALSSSFLFLYFNDRITSLIYVFOEFTKG
LMSPEFSYOLRCCSAKCYIQALITISIAVISTILFSGISGAAFTYLGWTFEIVAG
FCFIPAVIIRIYLPYVYSS"
5303..5402
/feature="REP01, repeated element"
complement(5424..6668)
```

```
CDS
/feature="rhlE"
/feature="RC1020"
complement(5424..6668)
/feature="rhlE"
/codon_start=1
/translation="ATP-dependent RNA helicase RhlE"
/product="ATP-dependent RNA helicase RhlE"
/protein_id="AAL03558.1"
/db_xref="GI:15620136"
/translation="MDYTKNKLSEELIATETNITTEPETEIOKOSIPYAMGSDIL
ASSQSGSKTALVLLPLDSFTKNKTALIVPEFRELATQISTLNKVTYSKINSAY
LIGEPMPKQFIQLKNPKVYIIGTPRIIDLHNRGSLKIDRIGTVLDEMDRLDMGM
KEOLEINKFLPEKROVLMESATMPKHIIAIVSOQLNPNPVITGVATNKAAAEIKOS
MHVSDKEKFSALTQOLNREGSVIIFVKTRASDQALAKMYENHTEALIHGDSORO
REVRTLSFRKSNHRIMVAVDVAARGIDPHOTOIVINDLPMCPEDYLAHRTGRGAC
TGHALSTSPDVTIRRAIDRLVYNKESIPRSFPRDKNNRKSFGKRAAGDGKRFNA
FDSKRRTFDGNSYGNKKRV"
complement(6808..6836)
/feature="REP05, Rickettsia pallidromic element 1"
complement(7165..7377)
/feature="cspa"
/feature="RC1021"
complement(7165..7377)
/feature="cspa"
/codon_start=1
/translation="cold shock-like protein"
/product="cold shock-like protein"
/protein_id="AAL03559.1"
/db_xref="GI:15620137"
/translation="MATNIVGKWKYNSKTNKFGFTIODNGKQDFVHKSAVDAAGLHS
LEEGDVFDELEKOGKAYAVNLRK"
7756..7838
/feature="RCRNA30"
7756..7838
/feature="RCRNA30"
/product="trna-Leu"
/feature="trna-Leu"
BASE COUNT 2993 a 1445 c 1263 g 2627 t
ORIGIN
Alignment Scores:
Pred. No.: 2,99e-38 Length: 8328
Score: 524.00 Matches: 100
Percent Similarity: 77.53% Conservative: 38
Best Local Similarity: 56.18% Mismatches: 40
Query Match: 59.21% Indels: 0
Gaps: 0
DB: 1
US-09-770-509-2 (1-178) x AE008653 (1-8328)
QY 1 AAlaserGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGly 20
Db 305 AGTCTAATCTGCAAGGCTCTAATTTGTAGTACGTATGACATGCAATGCTTGA 364
QY 21 ArgSerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAla 40
Db 365 CATCTTTATGCACTAACAATAATACACACGCGGTTCCTACGACTAGGCTTGGGCA 424
QY 41 GlySerLysProGlnLeuGlyLysArgSerAlaGlnGlnLysValAspIleGlnArg 60
Db 425 GGAGCTTCTCTGAGGTTGGACACTTGTCTCAAGAAACAAACGAAATTCGTAAT 484
QY 61 MetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyThrCysThr 80
Db 485 TACCTAGAAATATGCAATATGATATTATTACGCGAGGTATGGCGGTGATACGATACC 544
QY 81 GlyAlaAlaProValAlaAlaSerValAlaArgGlnLeuGlyIleLeuThrValGlyVal 100
Db 545 GGTCTCGACCGCATTTATGCAAGCATTAAGAACTAGACATCTTACGTTGGGGA 604
QY 101 ValSerThrProPheArgSerLeuGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
Db 605 GTACTAATACCTTTTCCATTGTAAGCGGTCATCGATGAAACCTAGCCATTAAGGACTT 664
```


LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
Db	22 SerLeuAlaProHisIstIstIeThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41	AF067822	1	GI:3426307		
Db	343 TCATTATTCACAGAAGAAGATCCCACTGGGATGGAACCTCCACAAAGCTTTGGGCGCTGT 402	AF067822	1	GI:3426307		
Qy	42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61	AF067822	1	GI:3426307		
Db	403 TCTCTCCCTGAGATGTCCTGCGCGCGCTGAGGAAATCAATAGAGAGATAAATCGCAGAG 462	AF067822	1	GI:3426307		
Qy	62 LeuGlnAspSerSerMetLeuPheIleThrGlyMetGlyGlyThrCysThrGly 81	AF067822	1	GI:3426307		
Db	463 ATCTCAGATGACATATGCTATTATTAATCTCGGATATGGGTGGAGGCACTGGAACTGGT 522	AF067822	1	GI:3426307		
Qy	82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValAl 101	AF067822	1	GI:3426307		
Db	523 GGAGCGCCGTATTTCGCTAAGCGTCGCAAAACACAGCAAAATTCCTAACTGTGGTGTGTC 582	AF067822	1	GI:3426307		
Qy	102 SerThrProPheArgSerGlyGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121	AF067822	1	GI:3426307		
Db	583 ACTAGGCCATTTTCATTTCGAAAGGCCCCACAGGATGAAACACGCGAAATACGCTCGGAG 642	AF067822	1	GI:3426307		
Qy	122 GlnLeuAlaLysThrValAspThrLeuIleValValProAsnGlnAsnLeuAlaLeu 141	AF067822	1	GI:3426307		
Db	643 GAGCGTCAAAAGCATGTGTGATACCTTATCTATTAATCTTAACAAATTTGTTGCTATA 702	AF067822	1	GI:3426307		
Qy	142 AlaAspLysSerThrThrMetLeuGluAlaPheArgThrAlaAspValLeuGlu 161	AF067822	1	GI:3426307		
Db	703 GCTAAGCAACACTACTTTCGCGTACAGCGCTTCAAGTTGCGCGACACTGTGTGCATACA 762	AF067822	1	GI:3426307		
Qy	162 GlyValLysGlyValIleThrAspLeuIleValArgProGlyLeuIleAsnLeu 178	AF067822	1	GI:3426307		
Db	763 GCGGTGCGGTATATCACAGACTTAATGTTATGCCAGGTTGATTAACCTA 813	AF067822	1	GI:3426307		
RESULT 8						
LOCUS	AF067822	1202 bp	DNA	linear	BCT 18-AUG-1998	
DEFINITION	Clostridium lentocellum cell division protein (ftsZ) gene, complete cds.					
ACCESSION	AF067822					
VERSION	AF067822.1	GI:3426307				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
MEDLINE						
PUBMED						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
MEDLINE						
PUBMED						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
MEDLINE						
PUBMED						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
MEDLINE						
PUBMED						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
MEDLINE						
PUBMED						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
MEDLINE						
PUBMED						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
MEDLINE						
PUBMED						
REFERENCE						
AUTHORS						


```

/db_xref="GI:3426308"
/translation="MLETWENLGAQIKVIGVGGGNNADVDMIEKLEGVEITVN
TDHQLARSGAPAKIOTIEKMTKRGLAGANPEIGTSAESREIITAIAGMLFTT
AGMGGTGTGAAPVIAIAKEBGLTVGVVTPESFEGRKRMNAKGIJLEKONVPT
LVVIPNDKILIOVLDKKTMMVDAFSKADVOGVGITPLISPGIINDPFAVRTM
NNKGVAMHGIRATGENRAREVKAISPLDTSIDGARCVLYNNCGGSLGMEAN
VNGKLREAVDPDAEILIGTSTINENGEIITIVATDQNHDSINTFKPVATKEE
TPVQTQATTAHVEETKPTPKIDQVEIIPQFLRRNR"
BASE COUNT      428 a 167 c 272 g 335 t
ORIGIN
Alignment Scores:
Pred. No.:      7.55e-39      Length:      1202
Score:          520.00      Matches:      103
Percent Similarity: 76.00%      Conservative: 30
Best Local Similarity: 58.86%      Mismatches: 42
Query Match:    58.76%      Indels:      0
DB:             1          Gaps:      0
US-09-770-509-2 (1-178) x AF067822 (1-1202)
QY      4  LeuGlUgLyVaLgUphElLeVaLAlAsnThrAspCySgLnAlaLeuGlYArGserLeu 23
Db      147 CTTCGAGCTGTGTAATTATTACTCTTAATACAGATCATCAAGCACTGGCTAGATCAGGT 206
QY      24  AlAProHIsLySleThrLeuGlYLySAsPleThrLySgLyLeuGlYAlaGlySerLyS 43
Db      207 GCACACCAACCAAAATTCAGATTGTGTGAAAAATAGACACGTCGAGCTTGAGAGCGCCTAAT 266
QY      44  ProGluLeuGlYLySArGserAlaGluGlnGlnLyVaLAsPleGlnArGmetLeuGln 63
Db      267 CCAGAAATGGTACAAAATACAGACAGAGAACTCGTAGACATCACTAAGCTATTTAAA 326
QY      64  AspSerAsnMetLeuPheIleThrGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
Db      327 GCGGCAATATGCTATTATTATACACCGCAATGGGTGGTGTGACAGTACAGTGCAGCA 386
QY      84  ProValValAlaSerValAlaArGgluLeuGlyIleLeuThrValGlyValSerThr 103
Db      387 CCAGTATATGCGAGTATTCGAAAGAGAGATTTTAAAGTTGTGTGTCGACAGCAAG 446
QY      104  ProPheArGserGluGlyProAsnArGthrArGleAlaAlaSnAlaGlyValLySgLeu 123
Db      447 CCATTATGTTTGAAGGTGCGCAAAAGATGATTAATGCGAGAAAGCTATTGCGAGAGCTT 506
QY      124  AlAlaLySerValaAsPthrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsP 143
Db      507 AACCAAAATGTAGATACCTTAGTTGTTATCCCAATGATAAATTTACAGATTATGAT 566
QY      144  LySerThrThrMetLeuGluAlaPheArGtyrAlaAsPAsPValLeuLeuGlnGlyVal 163
Db      567 AACAAAGCAACGAGGTGGACGCAATTTAGTAAGCGCCAGCATGATTACAAACAAAGGTGA 626
QY      164  LySgLyValThrAsPleuIleValArProGlyLeuIleAsnLeu 178
Db      627 CAGGGGATTACAGATTAAITTTCAACCCAGGATTATTAACTT 671
RESULT 9
AE013119/c      11683 bp      DNA      linear      BCT 09-MAY-2002
LOCUS      Thermoaerobacter tengcongensis strain MB4T, section 146 of 244 of
DEFINITION      the complete genome.
ACCESSION      AE013119 AE008691
VERSION      AE013119.1 GI:20516647
KEYWORDS
SOURCE      Thermoaerobacter tengcongensis.
ORGANISM      Thermoaerobacteriaceae; Thermoaerobacter.
REFERENCE
AUTHORS      Bao, Q., Tian, Y., Li, M., Xu, Z., Xuan, Z., Hu, S., Dong, W., Yang, J.,
Chen, Y., Xue, Y., Xu, Y., Lai, X., Huang, L., Dong, X., Ma, Y., Ling, L.,
Tan, H., Chen, R., Wang, J., Yu, J. and Yang, H.

```

```

TITLE      A Complete Sequence of the T. tengcongensis Genome
JOURNAL      Genome Res. 12 (5), 689-700 (2002)
MEDLINE      21992816
PUBMED      11997336
REFERENCE
AUTHORS      Bao, Q., Xu, Z., Hu, S., Dong, W., Chen, Y., Wang, J., Yu, J. and Yang, H.
TITLE      Direct Submissions
JOURNAL      Submitted (07-OCT-2001) Beijing Genomics Institute/Genomics and
Bioinformatics Center, Institute of Genetics and Development,
Chinese Academy of Sciences, Beijing Airport Industrial Zone B6,
Beijing 101300, China
3 (bases 1 to 11683)
REFERENCE
AUTHORS      Li, M., Xuan, Z., Yang, J., Ling, L. and Chen, R.
TITLE      Direct Submissions
JOURNAL      Submitted (07-OCT-2001) Institute of Biophysics, Chinese Academy of
Sciences, Beijing 100101, China
4 (bases 1 to 11683)
REFERENCE
AUTHORS      Tian, Y., Xue, Y., Xu, Y., Lai, X., Huang, L., Dong, X., Ma, Y. and Tan, H.
TITLE      Submitted
JOURNAL      Submitted (07-OCT-2001) Institute of Microbiology, Chinese Academy
of Sciences, Beijing 100080, China
FEATURES
source      1..11683
location/Qualifiers
1..11683
/organism="Thermoaerobacter tengcongensis"
/strain="MB4T"
/db_xref="taxon:119072"
/complement(88..107)
/note="putative rho-independent transcription terminator"
gene
/complement(109..1195)
/gene="ftsZ"
/note="TPE1639"
/complement(109..1182)
/gene="ftsZ"
/note="Best Blastp hit - g1101751781db1|BAB06277.1|
(AP001515) cell-division initiation protein (septum
formation) [Bacillus halodurans], score 352, E-value
4.00E-96"
/codon_start=1
/transl_table=11
/product="Cell division GTPase"
/protein_id="BAM24841.1"
/db_xref="GI:20516648"
/translation="MIGLETDVQFAIRKIVGVGGGNNVNRVAVGVKVEFTAIN
TDKQALSLSKAEKTIQIGELTKLGAGANPEIGTSAESREIERIKGADMFIT
AGMGGTGTGAAPVIAIAKEBGLTVGVVTPESFEGRKRMNAKGIJLEKQVDA
LITINDRLIOVEKETSMDKADLDVROGVGSDIIVAGLVAVDPADVTIM
VWTGIAHMGIGTASGKNATKAKQKAIHSLRSTSGSKGILINAGPMLTTFEVN
EAMNFIENADPDANITFGAVIDEALDDQIRITVATGFERNKAKEPKKREKVGVD
VKLEIDLDNDDDIDPFLRRGR"
/complement(526..1011)
/gene="ftsZ"
/note="Pfam match to entry tubulin, Tubulin/FtsZ family,
score 95.9, E-value 8.10E-25"
/complement(1191..1195)
/gene="ftsZ"
/complement(1197..2437)
/gene="ftsA2"
/note="TPE1640"
/complement(1197..2423)
/gene="ftsA2"
/note="Best Blastp hit - g11205671sp1|P28264|FTSA_BACSU
CELL DIVISION PROTEIN FTSA g11216939|pir1139847
cell-division protein (septum formation) ftsA - Bacillus
subtilis g11429401|gb|AA22456.1| (M22630) ftsA [Bacillus
subtilis] g112633899|emb|CAB13401.1| (289111)
cell-division protein [Bacillus subtilis], score 306,
E-value 2.00E-82"
/codon_start=1
/transl_table=11
/product="predicted Arpases of the HSP70 class involved in
cell division"
/protein_id="BAM24842.1"
/db_xref="GI:20516649"

```


Db 780 CCCTTCCTTTGAAGAGAGAGAAATGCGCAGCAGAGATGGCAATAGAAATTTA 721
Oy 124 Alalstytvalaapthrlleuilevalproaanglaasleuileualaasp 143
Db 720 AAGACAGACATGATGCTATACGATCCCTAATGATAGATTATACGAGTAGGAG 661
Oy 144 lysSerThrThrMetLeuGlualAphaeRgYrYalaaSPaSVAlleuGluglyVal 163
Db 660 AAGACAGACATGATGCTATGCTTTAAAGTGGCAGATGAGCTTTTAAGCAGGCGGT 601
Oy 164 lylgilyalThrappLeuilevalarProglyleuileasleu 178
Db 600 CAGGGAATTCGTATTATAGCTGTCCGGGGCTTGTGAATGTG 556
RESULT 10
RPX03
LOCUS 279110 bp DNA linear BCT 11-NOV-1998
DEFINITION Rickettsia prowazekii strain Madrid E, complete genome; segment 3/4.
ACCESSION AJ235272
VERSION AJ235272.1 GI:3861033
KEYWORDS complete genome.
SOURCE Rickettsia prowazekii.
ORGANISM Bacteria; Proteobacteria; alpha subdivision; Rickettsiales; Rickettsiaceae; Rickettsieae; Rickettsia; typhus group.
REFERENCE 1 (bases 1 to 279110)
AUTHORS Sacheritz-S.G., Zomorodipour,A., Andersson,J.O., Eriksson,A.S., Winkler,H.H. and Kurland,C.G.
TITLE The genome sequence of Rickettsia prowazekii and the origin of mitochondria
JOURNAL Nature 396 (6707), 133-140 (1998)
MEDLINE 99039499
PUBMED 9823893
REFERENCE 2 (bases 1 to 279110)
AUTHORS Andersson,S.G.E.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-1998) S.G.E. Andersson, Siv. Andersson@molbio.uu.se, Dept. of Molecular Biology, University of Uppsala, Husargatan 3, Uppsala, S-751 24, SWEDEN
FEATURES
source Location/Qualifiers
1. 279110
/organism="Rickettsia prowazekii"
/strain="Madrid E"
/db_xref="taxon:782"
63. 1046
/gene="Rp479"
63. 1046
/gene="Rp479"
/codon_start=1
/transl_table=1
/product="OCTAPRENYL-DIPHOSPHATE SYNTHASE (1spB)"
/protein_id="CA14934.1"
/db_xref="GI:3861034"
/translation="MNIIVKIQQLNKLDEVTQQLNLIISLKSDAELIEKYGKLVENG
GKRRLPLITITAKMEFDYKNNHKLASAVEIHAATLHADVNDSTLRFPETAV
TIGSKTSIIIVDPLFSOSFKLVASGCIKAMVNLKASVITSGEYVQVLKMERIT
TIDYQOIVAKSTALREAGACEVGAIIAQQVRAVSKDVQNFGLTITROYTDLIDY
LGSQVQKNTIGDPLBSGKVTLPFLVHLKLEQDQVLMENMLKSDKRTDVKLRD
LMKHAIVETVNTLSLENNANLNLRIPVNIKYLYFSIIRFLIRSY"
1065. 1138
/gene="tRNA Arg (CCG)"
1065. 1138
/gene="tRNA Arg (CCG)"
/gene="tRNA Arg (CCG)"
complement(1194. 1265)
/gene="tRNA Gln (TTG)"
complement(1194. 1265)
/gene="tRNA Gln (TTG)"
complement(1316. 3091)
/gene="Rp482"
complement(1316. 3091)
/gene="Rp482"
CDS

gene /codon_start=1
/transl_table=1
/product="unknown"
/protein_id="CA14935.1"
/db_xref="GI:3861035"
/translation="MRHDIRINLRLKLEIEYNIIGYIIPSDKNYNEVPEYAKRLEYI
TGFTSGSJAIIICKDAFFETDGRVLEQANKEDLAFYKLYDLKIEYTLDNKIGY
DPLFTYQVLANININENHINENLVKIVNKLPLEPSKLYLHDIKFAGVSNDRIR
CRETILSSSVGCMKNDNDIIVTIDASICIMLNLRASPVNTPLMEFAKVVITLSTK
LYLFINPRIDPEIINARPEITILPEKEFNILRDSKNRILIDSTISVHIMDLANK
KYKIVPECLLAKCNDIEIKHAIDPHINDAVALCFEAEFLYHSEENVNSCHSH
LIDESHSLCKLTAQAKOGEVYDSFHAICGQENSAIIHYRANPRAKRKEGHL
LIDSGAQYKGAATDITRIIVIGIPTCEQKRRYOVLGIALTRAKRPNKVTGNMID
ILARQYLMODMIDYDPHTGHWGVSFLSHEGPOSINSNKLILAKMILSNKPGYIP
GKVIYRIENLIVYKNNGLMEFEFLSLVPAKRLIDVALINDININIKKEYNKRIR
IYNLSTQAKVNLNRYNINFLQSL"
complement(3174. 4463)
/gene="Rp483"
complement(3174. 4463)
/gene="Rp483"
/codon_start=1
/transl_table=1
/product="PUTRESCINE-ORNITHINE ANTIporter (pole)"
/protein_id="CA14936.1"
/db_xref="GI:3861036"
/translation="MSOKLSFMAVFAVLTGSGIQTSVFIIPLSLAPGCVSIMGWILS
LFGMSALVPSGICICFPKPTGGPHVVRASPDPTIAPFGMTWITSFVSTIVIS
AIGVLPFRKSGTIIIDLIQLILALAIINLKPRAKYEPIITLKPPLVYGL
NALPHININVIYAKVENFTIPTMGRVALLTFWGIGIECATTTAGTVDPAATIP
RAIIIGTCVNAFLYIINSIGIILIPASELINSAPADATLFGGTWSVTVIAS
VICIETLNMVNIISGOIALGLAFLPKFPKNSNNAPIYGIISLGTPLLEF
SNNFPAKQITQIIDFVIAFLVYLCSLAFKLVIFSKENFSYLYLEVALIISIECT
WIVYKPEFELIACSETLIGIPVYGYWFLMNRKI"
complement(4633. 4965)
/gene="Rp484"
complement(4633. 4965)
/gene="Rp484"
/codon_start=1
/transl_table=1
/product="HESB PROTEIN (hesB2)"
/protein_id="CA14937.1"
/db_xref="GI:3861037"
/translation="MRNVISILTDPAAKOVALLIEKRAPTFGIRVGIKSGCAGCTY
VYADNKNQFDEVEVEKGVILIDPLTMYILASEMDYVETNFKSGFTTPNNEKANC
GGKSPSY"
complement(4969. 5364)
/gene="Rp485"
complement(4969. 5364)
/gene="Rp485"
/codon_start=1
/transl_table=1
/product="NIFU PROTEIN (nifU)"
/protein_id="CA14938.1"
/db_xref="GI:3861038"
/translation="MAYSKKVIDHYENPRNNGVSLDKKKKNVGTGLVGAPAGDVAKLQ
LEVGDDEITDAKFKFGGSAJASSLVYEMWIKGKSVEDAKEIKNFEIAKELSLPV
KHGSLADAIKAATADYKORENKKS"
complement(5430. 6662)
/gene="Rp486"
complement(5430. 6662)
/gene="Rp486"
/codon_start=1
/transl_table=1
/product="NIFS PROTEIN HOMOLOG (nifH)"
/protein_id="CA14939.1"
/db_xref="GI:3861039"
/translation="MNQQLKNTLPLPYMQSTTPIDPRVEMALPYFTTKGNPNRSH
SHSFGWEAENAVENARSNAKVLGADSKELIFSGATESNNLVYKGIAPYGNKKHI
ITLVSEHKCVLANCRHLEQGIKITVYIPISNGIIDETLKNATDQTLFYSVAVVN
ETGVIOPLKEIGKICRERNVFPFSDIAGCGKIPIDNENCDIASISGHIYGVK
GALYTRKPRVRYTPLINGGQGRGMRSGTLPPLIVGLGIASIAINEMKRDQVHN
TLDRPLNNHSHKISIEVYLVGDKDQDKGNLNSFAGVESITLAIKIDLVASSAC
TSASLEPSYVLSIGISELIAHNSIRGIRFTTEIDIVAVNIVCSKIDKLRLRLSLP

CDS

CDS	BASE COUNT
<pre> /gene="ftaz" <1..>697 /gene="ftaz" /codon_start=3 /transl_table=11 /protein_id="AAC44314.1" /db_xref="GI:1514678" /translation="AVNNMIQSNLCQVNFVANTDAQLEKSLCDKKIQLGINTLTKGL GAGLAPVPGKGAABESIDELMEHKDSHMLFTAGMGSGTGCAPIYAKNARARAV VKDGAKEREKILIVGVYTRKPFEGCVNRHMLAEGLLELDKYDTLIVIPNQLTGA NEKTTXYDAFQADLVNLHIGIRGVTDLMIIMPGLINLFDIETVMSMGKAMIGTGEA EGEGRASIAEAA" </pre>	<p>228 a 104 c 180 g 183 t 2 others</p>

Alignment Scores:	
Pred. No.:	7.62e-39
Score:	517.00
Percent Similarity:	72.49%
Best Local Similarity:	56.08%
Query Match:	58.42%
DB:	1
	Gaps: 1


```

/protein_id="AAB54067.1"
/db_xref="GI:2078543"
/translation="OS=NeoglyphisnotataLEKSLDCKKIOGINTKLGALP
DICKGAESIDEIMEHTRDMSHLEFTAGCGGTGTAAPYIAKAREARVAKDKA
KERKILIVGVVTRPGEFEGVRKMTAELELQYVDTLIVPNOLRIANETTF
ADAFOLADNVLIHIGRVTDLIMPGLINDFADIEVMSKRAMGTGEEDRA
ISAEALISNPLDNNVSMKGAOGLILNITGGDMTLEVDAAANRVEEDENANITF
GATPDVMEGRVR"
BASE COUNT      278 a      121 c      234 g      235 t
ORIGIN
Alignment Scores:
Pred. No.:      1,85e-38      Length:      868
Score:          514.00      Matches:      106
Percent Similarity: 71.43%      Conservative: 29
Best Local Similarity: 58.08%      Mismatches: 42
Query Match:    58.08%      Indels:      12
DB:             1      Gaps:      1
US-09-770-509-2 (1-178) x WSU95749 (1-868)
OY      2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
    ||| |||:::||||| |||:::||||| |||:::||||| |||:::||||| |||:::|||||
DB      5 TCCAAATTTCAGAGAGTAATTTGTTGTAGCAATACCGATGCTCGAGCGTTAGAGAG 64
OY      22 SerLeuAlaProHisIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      65 TCATTATGCGATAAAAAAATTTCAGCTTGATTAACCTTAAGAGCGCTTGCGCCGGA 124
OY      42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
    ||| |||:::||||| |||:::||||| |||:::||||| |||:::||||| |||:::|||||
DB      125 GCTTTCCTCGATATGTTGTAAGGTGCGAGCAAGAAATCAATTGATTAATGAGCAT 184
OY      62 LeuGlnAspSerAsnMetLeuPheIleThrGlyMetGlyGlyLysGlyThrCysThrGly 81
    ||| |||:::||||| |||:::||||| |||:::||||| |||:::||||| |||:::|||||
DB      185 ATTAGGACAGCCATATGCTCTTATCAGCAGGAGATGGGTGGTACTGGAACAGT 244
OY      82 AlaAlaProValAlaSerValAlaArgGluLeu----- 93
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      245 GCTGCACCGGTGATTCAGAAAGCAGCAAGAAAGCAGCGGTGTTAAGATTAAGGA 304
OY      94 -----GlyIleLeuThrValGlyValLysSerThrProPheArgSerGluGly 109
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      305 GCAAAAGAAAAAAGATCTGCTGAGTGTGAGTGAAGCGTTGCGTTTGAAGGT 364
OY      110 ProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysThrValAspThr 129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      365 GTGCGACGATGCGCACTGCAAGAGCTTGAGCTTGAAAGAGTTGCAAAATACGTAGATCA 424
OY      130 LeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaLysSerThrThrMetLeu 149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      425 CTATATGTCATTCCTCAATCAAAATTTATTTAGATTCTTAACGAGCAAAACATGTTGCT 484
OY      150 GluAlaPheArgThrAlaAspValLeuLeuGluGlyValLysGlyValThrAspLeu 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      485 GATGCAATTCACCTCCGCAATATGTTGCAATATGGCATTAAGAGAGTAACGATTTG 544
OY      170 IleValArgProGlyLeuIleAsnLeu 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      545 ATGATCATGCCAGACTGATTAATCTT 571
RESULT 15
WSU95750      868 bp      DNA      linear      BCT 11-AUG-1999
LOCUS      Wolbachia sp. 1148 cell division protein ftsz (ftsZ) gene, partial
DEFINITION      cds.
ACCESSION      U95750
VERSION      U95750.1 GI:2078544
KEYWORDS
SOURCE      Wolbachia sp. 1148.
ORGANISM      Wolbachia sp. 1148.
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Wolbachiae; Wolbachia.

```

```

REFERENCE      1 (bases 1 to 868)
AUTHORS      Jager,C.R., Pintureau,B. and Heddi,A.
TITLE      Comparison between phylogenetic trees of some Trichogramma species
            and their Wolbachia endosymbionts
JOURNAL      Russ. Entomol. J. 7 (3-4), 163-168 (1998)
REFERENCE      2 (bases 1 to 868)
AUTHORS      Jager,C.R., Pintureau,B. and Heddi,A.
TITLE      Direct Submission
JOURNAL      Submitted (31-MAR-1997) Laboratoire de Biologie Appliquee,
            INSA-Lyon, 20, av A. Einstein, Villeurbanne 69621, France
FEATURES
source
    1. 868
    /organism="Wolbachia sp. 1148"
    /strain="1148"
    /specific_host="Trichogramma cordubensis"
    /db_xref="taxon:59579"
    <1..>868
    /gene="ftsZ"
    <1..>868
    /gene="ftsZ"
    /codon_start=2
    /transl_table=11
    /product="cell division protein ftsZ"
    /protein_id="AAB54068.1"
    /db_xref="GI:2078545"
    /translation="OS=NeoglyphisnotataLEKSLDCKKIOGINTKLGALP
DICKGAESIDEIMEHTRDMSHLEFTAGCGGTGTAAPYIAKAREARVAKDKA
KERKILIVGVVTRPGEFEGVRKMTAELELQYVDTLIVPNOLRIANETTF
ADAFOLADNVLIHIGRVTDLIMPGLINDFADIEVMSKRAMGTGEEDRA
ISAEALISNPLDNNVSMKGAOGLILNITGGDMTLEVDAAANRVEEDENANITF
GATPDVMEGRVR"
BASE COUNT      279 a      121 c      233 g      235 t
ORIGIN
Alignment Scores:
Pred. No.:      1,85e-38      Length:      868
Score:          514.00      Matches:      106
Percent Similarity: 71.43%      Conservative: 29
Best Local Similarity: 58.08%      Mismatches: 42
Query Match:    58.08%      Indels:      12
DB:             1      Gaps:      1
US-09-770-509-2 (1-178) x WSU95750 (1-868)
OY      2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
    ||| |||:::||||| |||:::||||| |||:::||||| |||:::||||| |||:::|||||
DB      5 TCCAAATTTCAGAGAGTAATTTGTTGTAGCAATACCGATGCTCGAGCGTTAGAGAG 64
OY      22 SerLeuAlaProHisIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      65 TCATTATGCGATAAAAAAATTTCAGCTTGATTAACCTTAAGAGCGCTTGCGCCGGA 124
OY      42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
    ||| |||:::||||| |||:::||||| |||:::||||| |||:::||||| |||:::|||||
DB      125 GCTTTCCTCGATATGTTGTAAGGTGCGAGCAAGAAATCAATTGATGAATTAAGACAT 184
OY      62 LeuGlnAspSerAsnMetLeuPheIleThrGlyMetGlyGlyLysGlyThrCysThrGly 81
    ||| |||:::||||| |||:::||||| |||:::||||| |||:::||||| |||:::|||||
DB      185 ATTAGGACAGCCATATGCTCTTATCAGCAGGAGATGGGTGGTACTGGAACAGT 244
OY      82 AlaAlaProValAlaSerValAlaArgGluLeu----- 93
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      245 GCTGCACCGGTGATTCAGAAAGCAGCAAGAAAGCAGCGGTGTTAAGATTAAGGA 304
OY      94 -----GlyIleLeuThrValGlyValLysSerThrProPheArgSerGluGly 109
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      305 GCAAAAGAAAAAAGATCTGCTGAGTGTGAGTGAAGCGCTTGCGTTTGAAGGT 364
OY      110 ProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysThrValAspThr 129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      365 GTGCGACGATGCGCACTGCAAGAGCTTGAGCTTGAAAGAGTTGCAAAATACGTAGATCA 424
OY      130 LeuIleValAlaProAsnGlnAsnLeuLeuAlaLeuAlaLysSerThrThrMetLeu 149

```

```
Db      425 CTTATTTGTCATTCCTCCATCAAAATTTATTTAGAAATTCCTAACGAGAAACTACATTGCT 484
OY      150 GUALAPHEARGTYRFLAASPAPVALLEULEUGLUGLYVALLYSGLYVALTHRASPLeu 169
Db      485 GATGCATTTCAACCTCGCGCATTAATGTTCTGCATATTGSCATTAAGAGAGATACTGATTGG 544
OY      170 ILEVALARGPROGLYLEUILLEASLeu 178
Db      545 ATGATCATGCCAGACTGATTAATCTT 571
```

Search completed: June 2, 2003, 09:39:19
Job time : 2508 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 2, 2003, 08:44:29 ; Search time 245 seconds
(without alignments)
1636.146 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASOLEGVEFIVANTDCALG.....LLEGVKVTVLIVRGLINL 178

Scoring table:

BLOSUM62	
Xgapex 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2185239 segs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+pn.model -DEV=x1h
-O=/cgn2.1/USPTO.spool/US09770509/runat_27052003_083647_18810/app.query.fasta.1.327
-DB=N_Genseq_101002 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09770509.cgnl.1.263.grnat.27052003.083647.18810 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEJURY -NEG_SCORES=0 -MAIT -DSBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6
-FEAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

1:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1980.DAT.*
2:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1981.DAT.*
3:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1982.DAT.*
4:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1983.DAT.*
5:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1984.DAT.*
6:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1985.DAT.*
7:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1987.DAT.*
8:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1987.DAT.*
9:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1988.DAT.*
10:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1989.DAT.*
11:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1990.DAT.*
12:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1991.DAT.*
13:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1992.DAT.*
14:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1993.DAT.*
15:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1994.DAT.*
16:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1995.DAT.*
17:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1996.DAT.*
18:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1997.DAT.*
19:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1998.DAT.*
20:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA1999.DAT.*
21:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA2000.DAT.*
22:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA2001.DAT.*
23:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA2001B.DAT.*
24:	/SID52/gcgdata/geneseq/geneseq-n-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	510	57.6	213251	24	AB067193	Listeria innocua c
2	510	57.6	2944528	24	AB03041	Listeria monococ g
3	510	57.6	3011208	24	AB069245	Listeria innocua D
4	509	57.5	950	24	ABK77373	Bacillus clausii g
5	506	57.2	111309	20	AA20250	Borrelia burgdorfe
6	506	56.7	910715	20	AA20248	Borrelia burgdorfe
7	502	55.9	1233	23	ABK72889	Bacillus lichenifo
8	495	55.9	1233	23	AA52905	Enterococcus faeca
9	495	55.9	1239	23	AA51272	Enterococcus faeca
10	485	55.9	32768	20	AA13065	Enterococcus faeca
11	483	55.7	236589	24	ABA90521	Genomic sequence o
12	485	54.8	1140	22	AAH52021	Mycobacterium tube
13	485	54.8	4403765	22	AA199683	Mycobacterium tube
14	485	54.8	4411529	22	AA199682	Mycobacterium tube
15	480	54.2	1172	23	AA552133	Staphylococcus aur
16	480	54.2	1173	21	AA393935	Staphylococcus aur
17	480	54.2	1173	21	AA49958	S. pneumoniae mono
18	480	54.2	1179	23	AA554988	Staphylococcus aur
19	480	54.2	1185	24	ABN92534	Staphylococcus epi
20	472	53.3	1008	19	AA655296	DNA of filamentous
21	472	53.3	1260	20	AA19269	DNA of filamentous
22	472	53.3	1260	21	AA249959	Streptococcus pneu
23	472	53.3	1260	23	AA555744	Monomeric filament
24	472	53.3	2702	19	AA655251	Streptococcus pneu
25	472	53.3	21706	19	AAV52169	DNA of filamentous
26	470	53.1	1326	22	AA673325	C glutamicum codin
27	470	53.1	3931	18	AAV74653	Staphylococcus aur
28	470	53.1	349980	22	AA685851	C glutamicum codin
29	467	52.8	1278	24	ABN66155	Streptococcus poly
30	467	52.8	2155561	24	ABN71527	Streptococcus poly
31	466	52.7	29112	23	AA559609	Propionibacterium
32	464	52.4	1185	23	AA554272	Pseudomonas aerugi
33	464	52.4	1317	24	ABN66156	Streptococcus poly
34	464	52.4	7143	19	AAV58282	S. prodigiosus poly
35	463	52.3	1611	21	AA447312	Type II ftsZ sequen
36	463	52.3	1611	21	AA447312	Arabidopsis chloro
37	463	52.3	1628	19	AAV05550	Buchnera sp. genom
38	461	52.1	640681	24	ABA92787	E. coli proliferat
39	452	51.1	1152	21	AA65913	Escherichia coli m
40	452	51.1	1152	21	AA39397	E. coli DNA for ce
41	452	51.1	1152	23	AA552289	FtsZ1 gene sequenc
42	452	51.1	1255	21	AA447315	DNA encoding E. co
43	447	50.5	1152	24	ABK11439	Haemophilus influe
44	447	50.5	1266	23	AA553458	Haemophilus influe
45	447	50.5	1830121	17	AAAT42063	Haemophilus influe

ALIGNMENTS

RESULT 1	AB067193	standard: DNA; 213251 BP.
ID	AB067193	
AC	AB067193:	
XX	29-AUG-2002 (first entry)	
DT		
DE	Listeria innocua contig DNA sequence #6.	
XX		
KW	Antibacterial; Listeria; food contamination; mutational analysis;	
KW	Infection; ds.	
XX		
OS	Listeria innocua.	
XX		
PN	MO200228891-A2.	
XX		
PD	11-APR-2002.	

[illegible][illegible]

```

Query Match: 57.63% Indels: 0
DB: 24 Gaps: 0
US-09-770-509-2 (1-178) x ABA03041 (1-2944528)

QY 4 LeuGlUGlYValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
DB 2113565 GTTCAAGAGATGATATTCCTCCGTTAATACAGACCTCAAGACCTTAAGCAAAA 2113506
QY 24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
DB 2113505 GCGAAGAACAAATTAACAAACGCGACAAATTAACGCGGTGTTAAGTGGCGGAGACTGTA 2113446
QY 44 ProGluLeuGlyLysArgSerAlaGluGlnGluLysValAspIleGlnArgMetLeuGln 63
DB 2113445 CCGGAATGTTGTAATAAAGCTGCAAGAAAGAACCCGCAACAAATTGAAGACCTTTAAAA 2113386
QY 64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyThrCysThrGlyAlaAla 83
DB 2113385 GCGCTGATATGATATTCGTAACGCTGGAATGGCGGCGGAGACTGGAATCGGGCTGCA 2113326
QY 84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
DB 2113325 CCTGTTATCGCTCAAAATCGCAAAAGAAATGGCGCTTAAACAGTAGGTGTTACTGCA 2113266
QY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
DB 2113265 CCATTTGGTTTGAAGACCAAAACGTAACCAACCTTAAGTGAACAGAAACCAATG 2113206
QY 124 AlaLysTyrValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaAsp 143
DB 2113205 AAGAAGACGGGTGATACGTTAATGTTATTCCTAATGACCGTTTACTCAAAATGTTGAT 2113146
QY 144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
DB 2113145 AAAAATACACCGATGCTGAAGCTTCCGTAAGACAGATATATGTTTACGTAAGGTGTA 2113086
QY 164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
DB 2113085 CAAGGATTTCTGATTTGATTCGCTGCTGCTTAACTTA 2113041

RESULT 3
AB069245/c
ID AB069245 standard; DNA: 3011208 BP.
XX
AC AB069245;
XX
DT 29-AUG-2002 (first entry)
XX
DE Listeria innocua DNA sequence #684.
XX
KW Antibacterial; Listeria; food contamination; mutational analysis;
XX
OS Listeria innocua.
XX
PN WO200228891-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001MO-FR03061.
XX
PR 04-OCT-2000; 2000FR-0012697.
XX
PA (INSP) INST PASTEUR.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Kunst F, Glaeser P.
XX
DR WPI; 2002-332479/37.
XX
PT New genomic sequences from Listeria species, useful for detection,
treatment and prevention of infection, also related polypeptides,

```

```

PT antibodies and modulators -
XX Claim 5; SEQ ID 2058; 180pp; French.
XX
CC The present invention relates to nucleic acid sequences
CC (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
CC and primers for identification and/or detection of Listeria (e.g. as
CC contaminants in foods, or mutational analysis) and for analysis of
CC gene expression. Proteins encoded by the nucleic acid sequences can be
CC used to screen for compounds that modulate gene expression, replication
CC and pathogenicity of Listeria (potential therapeutic agents), also for
CC treating infections by Listeria, and are useful as immunogens in
CC anti-Listeria vaccines.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 3011208 BP; 941651 A; 568176 C; 559189 G; 942192 T; 0 other;

Alignment Scores:
Pred. No.: 4 29e-42 Length: 3011208
Score: 510.00 Matches: 98
Percent Similarity: 76.00% Conservative: 35
Best Local Similarity: 56.00% Mismatches: 42
Query Match: 57.63% Indels: 0
DB: 24 Gaps: 0

US-09-770-509-2 (1-178) x AB069245 (1-3011208)

QY 4 LeuGlUGlYValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
DB 2161845 GTTCAAGAGATGATATTCCTCCGTTAATACAGACCTCAAGACCTTAAGCAAAA 2161786
QY 24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
DB 2161785 GCGAAGAACAAATTAACAAATCGGTACAAATTAACCGGTGTTAAGTGGCGGAGACTGTA 2161726
QY 44 ProGluLeuGlyLysArgSerAlaGluGlnGluLysValAspIleGlnArgMetLeuGln 63
DB 2161725 CCTGAATGTTGTAATAAAGCTGCAAGAAAGTCCGCAACAAATTGAAGAACTTTAAAA 2161666
QY 64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyThrCysThrGlyAlaAla 83
DB 2161665 GCGCTGATATGATATTCGTAACGCTGGAATGGCGGCGGAGACTGGAATCGGTGTA 2161606
QY 84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
DB 2161605 CCTGTTATCGCTCAAAATCGCAAAAGAAATGGCGCTTAAGTGGTGTGTTACACGA 2161546
QY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
DB 2161545 CCATTTGGTTTGAAGACCAAAACGTAACCAACCTTAAGTGAACAGAAACCAATG 2161486
QY 124 AlaLysTyrValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaAsp 143
DB 2161485 AAGAAGACGGGTGATACGTTAATGTTATTCCTAATGACCGTTTACTCAAAATGTTGAT 2161426
QY 144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
DB 2161425 AAAAATACCGCAATGCTTAAGCTTCCGTAAGACAGATATATGTTTACGTAAGGGGTA 2161366
QY 164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
DB 2161365 CAAGGATTTCTGATTTGATTCGCTGCTGCTTAACTTA 2161321

RESULT 4
ABK77373
ID ABK77373 standard; DNA: 950 BP.
XX
AC ABK77373;
XX
DT 13-AUG-2002 (first entry)
XX

```

```

DE Bacillus clausii genomic sequence tag (GST) #216.
XX
KW Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX
OS Bacillus clausii.
XX
PN MO200229113-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-US31437.
XX
PR 06-OCT-2000; 2000US-0680598.
XX PR 27-MAR-2001; 2001US-279526P.
XX
PA (NOVO ) NOVOZYMES BIOTECH INC.
XX (NOVO ) NOVOZYMES AS.
PI Berka R, Clausen IG;
DR WPI: 2002-416684/44.
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second
PT Bacillus cells, by using substrate containing Bacillus genomic
PT sequenced tag array -
XX
PS Claim 11; SEQ ID NO 4664; 200pp: English.
XX
CC The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive
CC follow-up characterisation is unnecessary, when one spot on an array
CC equals one gene or one open reading frame, since sequence information is
CC available. This sequence represents a genomic sequence tag (GST) used in
CC the method of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 950 BP; 277 A; 184 C; 262 G; 227 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1,45e-46 Length: 950
Score: 509.00 Matches: 101
Percent Similarity: 74.86% Conservative: 30
Best Local Similarity: 57.71% Mismatches: 44
Query Match: 57.51% Indels: 0
DB: 24 Gaps: 0
US-09-770-509-2 (1-178) x ABK77373 (1-950)
QY 4 LeungluclyalglapheleValAlaksnrhraspcyGlnAlaleugllyarSerIeu 23
Db 103 CTGCAGAGGTGGTATTATTCACAGTCACATACAGATGCACAAGCACTCATTTATCAA 162
QY 24 Alaprophilsilertleuglsaplierntlrsglyleugllylaaglyserlys 43

```

Db	163	GGGGAAAGAAAGTTCACATTGGCGGCAAACTTACTGGCGGCTTGGCGTGGCGCAAC	222
Qy	44	ProGluLeuEnglyLyArgSerAlaGluGlnGlnLyValAspIleGlnArgMetLeuGln	63
Db	223	CCGAGAAATAGGCAAGAAACGGCGAGAGGAAGCAAGCAAGTGAAGAAATCTTACA	282
Qy	64	AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyIleLeuThrCysThrGlyAla	83
Db	283	GGTTCTGATATGCTCTTATTTACTGGCGGAATGGAGGAGGAACAGCAGCCAGCAGCT	342
Qy	84	ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValIleSerThr	103
Db	343	CCTGCATATGCTGAAGTGGCAAAAGAGCTTGGCGCATTTACGGTCCGGTGTGGACACGC	402
Qy	104	ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLySGLuLeu	123
Db	403	CCTTTCTCTTTGAAAGGCCGTGAAGCGGCAAAACCAAGCATCTCTGGATTCGACGCTTA	462
Qy	124	AlaLySTyValAspThrLeuIleValProAsnGlnAsnLeuAlaLeuAlaAsp	143
Db	463	AAAGAAAAAGTGGACACATTTGATCGTGAATCCAAATGACCGCTTATGGAAATGGTGTAT	522
Qy	144	LySserThrThrMetLeuGlnAlaPheArgTyrrAlaAspAspValLeuGluGlyVal	163
Db	523	AAAAATACCGCATCTTGAACCGTCCGCCGAGCGGATTAAGTATTGGCCCAAGAGTT	582
Qy	164	LySGLyValThrAspLeuIleValArgProGlyLeuIleAsnLeu	178
Db	583	CAAGCATTTCCGATTAAATTCACACGCTGGGTGATCAACCTT	627
RESULT 5			
ID	AAAX20250	standard; DNA; 111309 BP.	
XX	AAAX20250;		
AC	AAAX20250;		
XX	04-MAY-1999	(first entry)	
DT			
XX	Borrelia burgdorferi polynucleotide sequence #3.		
DE			
XX	Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;		
KW	epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;		
KW	infection; diagnosis; characterisation; detection; ds.		
OS	Borrelia burgdorferi.		
XX			
PN	W09858943-A1.		
XX			
PD	30-DEC-1998.		
XX			
PF	18-JUN-1998;	98WO-US12764.	
XX			
PR	03-SEP-1997;	97US-0057483.	
PR	20-JUN-1997;	97US-0050359.	
PR	22-JUL-1997;	97US-0053344.	
PR	22-JUL-1997;	97US-0053377.	
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
PA	(MED-) MEDIMUNE INC.		
XX	Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;		
PI	White OR;		
XX			
DR	WPI: 1999-081217/07.		
XX			
XX			
PT	New isolated Borrelia burgdorferi nucleic acids - used to develop		
PT	products for the detection, diagnosis, characterisation, prevention		
PT	and therapy of infections, particularly Lyme disease		
XX			
PS	Claim 1; Page 738-800; 1128pp; English.		
CC	AAAX20248 to AAAX20402 represent polynucleotide sequences isolated from		
CC	Borrelia burgdorferi (Bb). Products derived from Bb can be used for		

CC the detection, diagnosis, characterisation, prevention and therapy of
CC Bb infections, e.g. Lyme disease. They can also be used for the
CC production of biosynthetic products, e.g. enzymes. Borrelia belongs
CC to a family of motile, spiral-shaped bacteria called Spirochetes.
CC Spirochetes are pathogenic in humans and Borrelia causes epidemic and
CC endemic relapsing fever, and Lyme borreliosis, more commonly known as
CC Lyme disease.

XX Sequence 111309 BP; 35956 A; 13151 C; 19075 G; 43117 T; 10 other;

Alignment Scores:

Score:	1.58e-43	Length:	111309
Best Local Similarity:	506.00	Matches:	101
Best Local Similarity:	74.42%	Conservative:	27
Query Match:	58.72%	Mismatches:	44
	57.18%	Indels:	0
	20	Gaps:	0

US-09-770-509-2 (1-178) x AAX20250 (1-111309)

OY 7 ValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaProHis 26

Db 12947 GTTGAATTTTGTGGCTAATACCGATCTTCAGGCTTCACAACTTCTATGCTCCCAT 13006

OY 27 LysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeu 46

Db 13007 AAAATGGCCCTGGAGCAAAAGATTACAGCAGCGCTTGGTGGGGAAGCCTGAGATT 13066

OY 47 GlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsn 66

Db 13067 GGACAGCTCAGCAGAGAGAACATAGATGTTAAACGAAATCATCTTCTGTCGCCGAT 13126

OY 67 MetLeuPheIleThrGlyLysMetGlyLysGlyThrCysThrGlyAlaAlaProValAla 86

Db 13127 ATGCTGTTTATTTACTGCTGCTGATGGGGCGGAGACAGAACCGGACACTCCAGTTAT 13186

OY 87 AlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPheArg 106

Db 13187 GCCCAAGTTCAGAAAGAGCTGTGATTTTAAACAGTTGAGCTGTAAACAAAGCCTTTAA 13246

OY 107 SerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyr 126

Db 13247 TTGGAAGTCTTAAGAGTGAAGTCTGTCGACGAGGAATTAATTAAGAAAGTCT 13306

OY 127 ValAspThrLeuIleValAlaProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThr 146

Db 13307 GTGATATCATGTGATCATTTCCAAATCAAAAGCTTTTAACGTGCTTACAAAGAAC 13366

OY 147 ThrMetLeuGlnAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyVal 166

Db 13367 ACCATTAAAGATGCTTTAAACGTCAGATGATGTTCTTAGAATGGCGCTTCAAGTTAT 13426

OY 167 ThrAspLeuIleValArgProGlyLeuIleAsnLeu 178

Db 13427 GCAGGCTTATTTATTTGAGCATGAGAGGTTAATATT 13462

RESULT 6 AAX20248 AAX20248 standard; DNA; 910715 BP.

XX AAX20248:

XX 04-MAY-1999 (first entry)

XX Borrelia burgdorferi polynucleotide sequence #1.

XX Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;

XX epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;

XX infection; diagnosis; characterisation; detection; ds.

XX OS Borrelia burgdorferi.

XX PN WO958943-A1.

XX 30-DEC-1998.

XX 18-JUN-1998; 98MO-US12764.

XX 03-SEP-1997; 97US-0057483.

XX 20-JUN-1997; 97US-0050359.

XX 22-JUL-1997; 97US-0053344.

XX 22-JUL-1997; 97US-0053377.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (MEDI-) MEDIMUNE INC.

XX Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;

XX White OR;

XX WPI: 1999-081217/07.

XX New isolated Borrelia burgdorferi nucleic acids - used to develop

XX products for the detection, diagnosis, characterisation, prevention

XX and therapy of infections, particularly Lyme disease

XX Claim 1; Page 157-671; 1128pp; English.

XX AAX20248 to AAX20402 represent polynucleotide sequences isolated from

XX Borrelia burgdorferi (Bb). Products derived from Bb can be used for

XX the detection, diagnosis, characterisation, prevention and therapy of

XX Bb infections, e.g. Lyme disease. They can also be used for the

XX production of biosynthetic products, e.g. enzymes. Borrelia belongs

XX to a family of motile, spiral-shaped bacteria called Spirochetes.

XX Spirochetes are pathogenic in humans and Borrelia causes epidemic and

XX endemic relapsing fever, and Lyme borreliosis, more commonly known as

XX Lyme disease.

XX Sequence 910715 BP; 327171 A; 129646 C; 130753 G; 323091 T; 54 other;

XX Alignment Scores:

XX Pred. No.: 2.47e-42 Length: 910715

XX Score: 506.00 Matches: 101

XX Percent Similarity: 74.42% Conservative: 27

XX Best Local Similarity: 58.72% Mismatches: 44

XX Query Match: 57.18% Indels: 0

XX DB: 20 Gaps: 0

XX US-09-770-509-2 (1-178) x AAX20248 (1-910715)

OY 7 ValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaProHis 26

Db 602333 GTTGAATTTTATTTGCTAATACCGATCTTCACAACTTCTATGCTCCCAT 602392

OY 27 LysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeu 46

Db 602393 AAAATGGCCCTGGAGCAAAAGATTACAGCAGCGCTTGGTGGGGAAGCCTGAGATT 602452

OY 47 GlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsn 66

Db 602453 GGACAGCTCAGCAGAGAGAACATAGATGTTAAACGATGCTTTCGAGCCAT 602512

OY 67 MetLeuPheIleThrGlyLysMetGlyLysGlyThrCysThrGlyAlaAlaProValAla 86

Db 602513 ATGCTGTTTATTTACTGCTGCTGATGGGGCGGAGACAGAACCGGACGCTCCAGTTAT 602572

OY 87 AlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPheArg 106

Db 602573 GCAGCAAGTTCAGAAAGAGCTGTGATTTTAAACGTGATGTAACAAAGCCTTTAAAG 602632

OY 107 SerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGlyValLysTyr 126

Db 602633 TTGGAAGTCTTAAGAGTGAAGTCTGTCGACGAGGATTAATACTTAAGAAAGTCT 602692

OY 127 ValAspThrLeuIleValAlaProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThr 146

Db 602693 GTAGATACATGATCATTTATTCACAAATCAAAAGCCTTTAACTGTTGACAAAGAAC 602752

QY 147 ThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuGluGlyValIysGlyVal 166
 DB 602753 ACCATTAAAGATCTTTAAAGCGGAGATGATCTTCTTAATAAGGGCTTCAAGGTATT 602812
 QY 167 ThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
 DB 602813 GCAGGGCTTATTATTCAGCATGGAGAGGTAAATATT 602848

RESULT 7
 ABK72889
 ID ABK72889 standard; DNA: 1134 BP.
 AC ABK72889;
 XX 13-APR-2002 (first entry)
 DE Bacillus licheniformis genomic sequence tag (GST) #180.
 KM Differential gene expression; genomic sequenced tag; GST;
 KW altered culture condition; environmental stress;
 KW physiological provocation; ds.
 XX Bacillus licheniformis.
 OS Bacillus licheniformis.
 PN WO200229113-A2.
 PD 11-APR-2002.
 XX 05-OCT-2001; 2001WO-US31437.
 PF 06-OCT-2000; 2000US-0680598.
 PR 27-MAR-2001; 2001US-279526P.
 XX (NOVO) NOVOZYMES BIOECH INC.
 PA (NOVO) NOVOZYMES AS.
 PI Berka R, Clausen IG;
 XX WPI: 2002-416684/44.
 DR Monitoring differential expression of several genes in first Bacillus
 PT cell relative to expression of same genes in one or more second
 PT Bacillus cells, by using substrate containing Bacillus genomic
 PT sequenced tag array -
 PS Claim 4; SEQ ID NO 180; 200pp; English.

CC The invention describes a method of monitoring differential expression of
 CC genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus
 CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive
 CC follow-up characterisation is unnecessary, when one spot on an array
 CC equals one gene or one open reading frame, since sequence information is
 CC available. This sequence represents a genomic sequence tag (GST) used in
 CC Note: The method of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 1134 BP; 333 A; 277 C; 307 G; 217 T; 0 other;

Alignment Scores:
 Pred. No.: 1,08e-45 Length: 1134
 Score: 502.00 Matches: 97
 Percent Similarity: 74.01% Conservative: 34
 Best Local Similarity: 54.80% Mismatches: 46
 Query Match: 56.72% Indels: 0
 DB: Gaps: 0

US-09-770-509-2 (1-178) x ABK72889 (1-1134)
 QY 2 SerGlnLeuGluGlyValAlaGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
 DB 97 AATGACGTTCAAGGAGGTGAGTTTATCCAGTCACACAGCAGATGCTCAGGCTCCAACTG 156
 QY 22 SerLeuAlaProHIsLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
 DB 157 TCMAAAGCGGAAAGCAAAATCGAGATCGGTGGAAGCTGACCGCGCGCTCGCGCGGA 216
 QY 42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
 DB 217 GCCAATCCGAGAGTGGGCAAAAACCCGACAGAGAAAGCAAAATTCGAAGAGCG 276
 QY 62 LeuGlnAspSerAsnMetLeuPheIleThrLysGlyMetGlyValGlyThrGly 81
 DB 277 CTGAAGGTGCAGATATGCTGTCACAGCCGGAATGGCGGCGGAGAACGAGCGGC 336
 QY 82 AlaAlaProValAlaAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValAla 101
 DB 337 GCGGACCTGTCATCGCACAAATCGCAAAAGATCGGGGATGATGATCGGCTGTC 396
 QY 102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValys 121
 DB 397 ACAAGGCCGTTTACCTTGAAGAGAGAAAGAACACCTCAGGCTCAGCGGTATTTC 456
 QY 122 GluLeuAlaIysTyrValAspThrLeuIleValAlaProAsnGlnAsnLeuLeuAlaLeu 141
 DB 457 GCATGAAGGAAGCGGTGACACCCGTATCGTCATTCGAAATGACCGTCTTCTTGAATC 516
 QY 142 AlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuGlu 161
 DB 517 GTCGATTAACAAACACACCGATCTTGAACCGTTCGTGAAGCGGACCAACCTTCGCCCA 576
 QY 162 GlyValLysGlyValAlaThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
 DB 577 GGTCTTACGAGCATTTTCAGACTGATCCCAACGCGCTGACGTATCAACTT 627

RESULT 8
 AAS52905
 ID AAS52905 standard; DNA: 1233 BP.
 AC AAS52905;
 XX 13-FEB-2002 (first entry)
 DE Enterococcus faecalis DNA for cellular proliferation protein #333.
 XX Antisense: ds; prokaryotic cellular proliferation gene;
 KW antibiotic; antibacterial; drug design.
 OS Enterococcus faecalis.
 XX WO200170955-A2.
 PN 27-SEP-2001.
 PD 21-MAR-2001; 2001WO-US09180.
 PF 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ
PI Yamamoto RT, Xu HH;
XX
XX
DR WPI: 2001-611495/70.
DR P-PSDB: AAU35046.
XX
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX
PS Claim 27; Seq ID No 6542; 51pp; English.

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antimicrobics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC [ftp.wipo.int/pub/publ/publ_seq_sequences](http://wipo.int/pub/publ/publ_seq_sequences).

SQ Sequence 1233 BP; 437 A; 223 C; 270 G; 303 T; 0 other;

Alignment Scores:	
Pred. No.:	7.06e-45
Score:	495.00
Percent Similarity:	73.1%
Best Local Similarity:	57.1%
Query Match:	55.9%
DB:	23
Length:	123
Matches:	100
Conservative:	28
Mismatches:	47
Indels:	0
Gaps:	0

US-09-770-509-2 (1-178) x AAS52905 (1-1233)

QY	4	leugluGlYValGluPheIleValIAsnThrAspCGlnAlaIleuGlyArgSerLeu	23
		106 GTTAAAGCGCGGAATTTATTCACACCCCAATACAGAGCTTCAACCAATTAACCAATTCACAAA	165
Db			
QY	24	AlAProHIslyIeThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys	43
		166 GCAGAAACAGTGATTCATTATAGGCCCTAAATACACCTCGTGTTAGTGGCGGTTCCAA	223
Db			
QY	44	ProGluLeuGlyLysArgSerAlaGluGlnGlyValAspIleGlnArgMetLeuGln	63
		226 CCTGAAGTGGCCCAAAAAGCTCAGAGAGAAAGTGAACAAGTATTTCAGAAATCATTAACA	285
Db			
QY	64	AspSerAsnMetLeuPheIleThrGlyMetGlyGlyIYThrCysThrGlyAlaAla	83
		286 GCGCGGATTAAGATTTTCAATTACTCGTATAGGGGCGGCAACTGTACAGTGGCGCG	343
Db			
QY	84	ProValValAlaSerValAlaArgGluLeuGlyIleThrValGlyValIAsnThr	103
		346 CCACTAGTGCAGAAAATCGCTAAAGATTAAGCGCTTAAACAGTTGGTGAATACCGT	405
Db			
QY	104	ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu	123
		406 CCAATTAGTTTGAAGGTCGCAAAACGTGGTGGTTTTGCCGCGGAAGCAATTGGCTTATTA	465
Db			
QY	124	AlAlaIstYrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp	143

Db	466	AAAGAAACGTTGATACACTATTATTATCTCAAAACCGCTATTAGAAAGCGTTGAC	522
		::: :::::	
Qy	144	LysSerThrThrMetLeuGlnAlaPheArgTrpAlaAspAspValLeuLeuGluGlyVal	164
Db	526	AAGAAAGAGCCCATCTTGAAACATTAAAGAAAGCTGATATATCTATTACTCAAGGTGT	586
Qy	164	LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu	178
		::: :::::	
Db	586	CAAGGATTTTCAGATTATATCTACTGCACCAAGGTACGTAATACCTTA	630

RESULT 9
AAS51272

AC	AAS51272;
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	Enterococcus faecalis DNA for cellular proliferation protein #49
XX	
KW	Antisense; ds; prokaryotic cellular proliferation gene;
KM	antibiotic; antibacterial; drug desig.
XX	
OS	Enterococcus faecalis.
XX	
PN	WO200170955-A2.
XX	
PD	27-SEP-2001.
XX	
PF	21-MAR-2001; 2001WO-US09180.

PA (ELIT-) ELITRA PHARM INC

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GT,
PI Yamamoto RT, Xu HH;

PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 27; Seq ID No 3854, 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC http://wipo.int/pub/published_pct_sequences.

Sequence 1239 BP; 439 A; 224 C; 273 G; 303 T; 0 other;

Alignment Scores:

Score:	7.11e-45	Length:	1239
Percent Similarity:	495.00	Matches:	100
Best Local Similarity:	73.14%	Conservative:	28
Query Match:	57.14%	Mismatches:	47
	55.93%	Indels:	0
		Gaps:	0

US-09-770-509-2 (1-178) x AAS51272 (1-1239)

```

OY      4 LeuGIuGIyValGIuPheIleValAlaAsnThrAspCysGlnAlaLeuGIyArgSerLeu 23
DB      106 GTTAAAGCGGTGAATTTATACAGCCAAATACAGCTTCAAGCATTTAAACATTCAAA 165
OY      24 AlaProHisIleThrLeuGIyLysAspIleThrIysGIyLeuGIyAlaGIySerIys 43
DB      166 GCAGAAACAGTGTATCAATTAAGCCCTAAATACCTGCTGTAGGTGCGGTCACAA 225
OY      44 ProGIuLeuGIyLysArgSerAlaGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 63
DB      226 CCTGAAGTGGCCAAAGAGCTGCAGAAAGAGAAAGAGAGATTCAGATTCATTCACA 285
OY      64 AspSerAsnMetLeuPheIleThrGIyMetGIyGIyGIyThrCysThrGIyAlaAla 83
DB      286 GCGGCGGATATGATTTTCACTTACTGCTGTATGCGTGTGCGGACACTGTACAGTGTCTCG 345
OY      84 ProValIAlaIAserValAlaArgGIuLeuGIyIleLeuThrValGIyValIAserThr 103
DB      346 CCAGTAGTTCGCAAAATCCCTAAGATTAAGGCGCTTAAACAGTGTGTGTACTGCT 405
OY      104 PropheArgSerGIuGIyProAsnArgThrArgLeuAlaAsnAlaGIyValIysGIuLeu 123
DB      406 CCATTAGTGTGAAGCTCCAAAGAGTGTGCTGCTGCGGTGAAGATTCGCTATATA 465
OY      124 AlaIysTyValAspThrLeuIleValIAlaProAsnGlnAsnLeuLeuAlaLeuAla 143
DB      466 AAGAAACGTTGATACCTATTATTAATTAATCTCAACACCGCTTATTAGAGTCTGTAC 525
OY      144 LysSerThrThrMetLeuGlnAlaIlePheArgTyValAlaAspValIleLeuGIuGIyVal 163
DB      526 AAGAAACGCCCAATGCTTGAAGCATTTAGAGAACGCTATATATATATATATATATATAT 585
OY      164 LysGIyValIThrAspLeuIleValIArgProGIyLeuIleAsnLeu 178
DB      586 CAAGGATTTCAAGATTTATCACTGCACACGATTACGTAACCTTA 630

RESULT 10
AAI13065/C
ID      AAI13065 standard; DNA; 32768 BP.
XX
XX      AAI13065;
AC
XX      19-MAR-1999 (first entry)
DT
XX
DE      Enterococcus faecalis genome contig SEQ ID NO:128.
XX
XX      Enterococcus faecalis; contig; detection; Enterococcal infection;
RW      vaccine; attenuation; computer readable medium; ds.
XX
XX      Enterococcus faecalis.
OS
XX      MO9850555-A2.
PN
XX      12-NOV-1998.
PD
XX      04-MAY-1998; 98MO-US08985.
PF
XX      14-NOV-1997; 97US-0066009.
PR      06-MAY-1997; 97US-0044031.
PR      16-MAY-1997; 97US-0046655.
XX

```

PA (HUMA-) HUMAN GENOME SCI INC.

PI Barash SC, Dillon FJ, Kunsch CA;

DR WPI; 1999-045171/04.

XX New isolated Enterococcus faecalis polynucleotides and polypeptides
 PT - used to develop products for the detection of Enterococcus and for
 PT use in vaccines for prevention or attenuation of Enterococcus
 PT infection.

PS Claim 1; Page 767-783; 2084pp; English.

XX A computer readable medium has been developed which has recorded on it
 CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
 CC AAI12938 to AAI13919 represent these nucleotide sequences which are
 CC primary nucleotide sequences, also known as contigs. The computer-based
 CC system can identify fragments of the Enterococcus faecalis genome with
 CC commercial importance. The products can be used to detect the presence
 CC of Enterococcus faecalis in samples. They can also be used for
 CC diagnosing Enterococcal infection in an animal and monitoring
 CC progression of disease, and for identifying agents which can be used to
 CC modulate the growth or pathogenicity of Enterococcus faecalis, or
 CC another related organism, in vivo or in vitro. In particular the
 CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences
 CC can be used in vaccines to prevent or attenuate an Enterococcal
 CC infection.

XX Sequence 32768 BP; 9529 A; 6883 C; 5611 G; 10720 T; 25 other;

Alignment Scores:

Pred. No.:	5.17e-43	Length:	32768
Score:	495.00	Matches:	100
Percent Similarity:	73.14%	Conservative:	28
Best Local Similarity:	57.14%	Mismatches:	47
Query Match:	55.93%	Indels:	0
		Gaps:	0

US-09-770-509-2 (1-178) x AAI13065 (1-32768)

```

OY      4 LeuGIuGIyValGIuPheIleValAlaAsnThrAspCysGlnAlaLeuGIyArgSerLeu 23
DB      27971 GTTAAAGCGGTGAATTTATACAGCCCAATACAGCTTCAAGCATTTAAACATTCAAA 27912
OY      24 AlaProHisIleThrLeuGIyLysAspIleThrIysGIyLeuGIyAlaGIySerIys 43
DB      27911 GCAGAAACAGATTCATTAATTAAGCCCTAAATACCTGCTGTAGGTGCGGTCACAA 27852
OY      44 ProGIuLeuGIyLysArgSerAlaGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 63
DB      27851 CCTGAAGTGGCCAAAGAGCTGCAGAAAGAGAAAGAGAGATTCAGATTCATTCACA 27792
OY      64 AspSerAsnMetLeuPheIleThrGIyMetGIyGIyGIyThrCysThrGIyAlaAla 83
DB      27791 GCGGCGGATATGATTTTCACTTACTGCTGTATGCTGCGGAGACTGTACAGTGTCTCG 27732
OY      84 ProValIAlaIAserValAlaArgGIuLeuGIyIleLeuThrValGIyValIAserThr 103
DB      27731 CCAGTAGTTCGCAAAATCCCTAAGATTAAGGCGCTTAAACAGTGTGTAGTACTGCT 27672
OY      104 PropheArgSerGIuGIyProAsnArgThrArgLeuAlaAsnAlaGIyValIysGIuLeu 123
DB      27671 CCATTAGTGTGAAGCTCCAAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 27612
OY      124 AlaIysTyValAspThrLeuIleValIAlaProAsnGlnAsnLeuLeuAlaLeuAla 143
DB      27611 AAGAAACGTTGATACCTATTATTAATTAATCTCAACACCGCTTATTAGAGTCTGTAC 27552
OY      144 LysSerThrThrMetLeuGlnAlaIlePheArgTyValAlaAspValIleLeuGIuGIyVal 163
DB      27551 AAGAAACGCCCAATGCTTGAAGCATTTAGAGAACGCTATATATATATATATATATAT 27492
OY      164 LysGIyValIThrAspLeuIleValIArgProGIyLeuIleAsnLeu 178

```


DB 27491 CAAGGATTTTCAGATTAACTGACGACGAGTACGTAACCTTG 27447
RESULT 11
ID ABA90521 standard; DNA: 2365589 BP.
XX ABA90521:
AC 16-MAY-2002 (first entry)
DT 16-MAY-2002 (first entry)
XX
DE Genomic sequence of *Lactococcus lactis* IL1403.
XX
KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese; ds.
XX
OS *Lactococcus lactis* IL1403.
XX
PN FR2807446-A1.
XX
PD 12-OCT-2001.
XX
PF 11-APR-2000; 2000FR-0004630.
XX
PR 11-APR-2000; 2000FR-0004630.
XX
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
XX
PI Bolotine A, Sorokline A, Renault P, Ehrlich SD;
XX
XX WPI: 2002-043418/06.
PT New nucleotide sequence useful in the identification or *Lactococcus*
PT lactis and related species -
XX
PS Claim 1; SEQ ID 1; 2504pp; French.
XX
CC The present invention is related to a *Lactococcus lactis* nucleotide
CC sequence (ABA90521) and related proteins (ABBS3300-ABBS5621). The
CC nucleic acid sequence is useful in the detection and/or amplification of
CC nucleic acid sequence, particularly to identify *Lactococcus lactis* or
CC related species. The proteins of the invention are useful for the
CC biosynthesis or biodegradation of a composition of interest. The
CC invention helps research in lactic bacteria, particularly useful in the
CC production of yogurt and cheese.
CC Note: The sequence data for this patent is based on equivalent patent
CC WO200177334 (published 18-OCT-2001) which is available in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pt_sequences.
XX
XX Sequence 2365589 BP; 765914 A; 415261 C; 420487 G; 763927 T; 0 other;
Alignment Scores:
Pred. No.: 2,32e-40 Length: 2365589
Score: 493.00 Matches: 98
Percent Similarity: 72.57% Conservative: 29
Best Local Similarity: 56.00% Mismatches: 48
Query Match: 55.71% Indels: 0
DB: 24 Gaps: 0
US-09-770-509-2 (1-178) x ABA90521 (1-2365589)
QY 4 LeuGluGlyValIgluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
DB 1938822 GTTCTGGCGTTGAATTTATCCGCCGTAATACGTGATGTACAGCATTTGCGTCAAAA 1938763
QY 24 AlaProHisLeuSerIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
DB 1938762 GCAGATACAGATTATTCACTGGGCCAAATTTGACTGCGGTTTGGGTGCTGTCGCCAA 1938703
QY 44 ProGluLeuGlyLysArgSerAlaGluGlnLysValAspIleGlnArgMetLeuGln 63
DB 1938702 CCTGAAGTGTGTAAGCGCGCAGCATGAGATACAGTCAACAGATACAGCCCTTGA 1938643
QY 64 AspSerAsnMetLeuPheIleThrGlyMetCylGlyLysThrCysThrGlyAlaAla 83

DB 1938642 GCCTCAGATATGATTTTCATCACTGCCGATGCGTGGTGGAGTCAAGTGTGCGCA 1938583
QY 84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValSerThr 103
DB 1938582 CCAGTTATCGCAACAAATTTGCTAAAGAGCTTGGCGCTTGTACTGTGTTACTCGT 1938523
QY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
DB 1938522 CCATTGGTTTGAAGATCAAAACGTTCTACTTTCACACTGAAGAAATTGAAGCTT 1938463
QY 124 AlaLysTyrValAspThrLeuIleValAlaProAsnGlnAsnLeuLeuAlaAsp 143
DB 1938462 CGTCAAAATGTTGATACACTTTCATATTTCAACAATACACTGCTTGAATTTGTTGAT 1938403
QY 144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspValLeuLeuGlyVal 163
DB 1938402 AAGAAACCTCCTTTAACTGAAAGCTTGGCGTAAGCTGATATGTTTAACTGCAAGCTGT 1938343
QY 164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
DB 1938342 CAAGGGTAACTGACTTGTACTTAACCCAGAAATGATTAACTT 1938298
RESULT 12
ID AAH52021 standard; DNA: 1140 BP.
XX
XX AAH52021:
AC AAH52021:
XX
DT 04-SEP-2001 (first entry)
XX
DE Mycobacterium tuberculosis potential drug target gene SEQ ID 75.
XX
KW Drug target; growth; organism viability; characterisation; ds.
XX
OS Mycobacterium tuberculosis.
XX
XX WO200135317-A1.
XX
XX 17-MAY-2001.
XX
PF 13-NOV-2000; 2000MO-US31152.
XX
PR 12-NOV-1999; 99US-0165086.
XX
PR 12-NOV-1999; 99US-0165124.
XX
PR 01-FEB-2000; 2000US-0179531.
XX
XX (BECG) UNIV CALIFORNIA.
XX
XX Eisenberg D, Rotstein SH, Marcotte EW;
XX
XX WPI: 2001-329193/34.
XX
XX P-PSDB; AAG81170.
XX
PT Identifying nucleotide or polypeptide sequence for use as drug target,
PT involves providing algorithm that analyzes a functional relationship
PT between nucleotide or polypeptide sequences, and comparing the
PT sequences -
XX
PS Disclosure; Page 97; 207pp; English.
XX
XX This invention relates to a method for identifying a nucleotide or
XX polypeptide sequence that may be a drug target, or essential for growth
XX or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092
XX represent DNA encoding proteins AAG81096 - AAG81241. Mycobacterium
XX tuberculosis proteins which are potential drug targets. The DNA and
XX protein sequences are used to illustrate the method of the invention. The
XX method involves providing an unknown nucleotide or polypeptide sequences,
XX and comparing it to a number of sequences along with at least one
XX algorithm capable of analyzing a functional relationship between
XX nucleotide and polypeptide sequences. The method is useful for
XX characterising the function of nucleic acids and polypeptides that may be
XX useful as a target for a drug or essential for the growth or viability of

CC an organism.
 XX
 SQ Sequence 1140 BP; 203 A; 368 C; 388 G; 181 T; 0 other.

Alignment Scores:

Pred. No.:	8.02e-44	Length:	1140
Score:	485.00	Matches:	95
Percent Similarity:	74.29%	Conservative:	35
Best Local Similarity:	54.29%	Mismatches:	45
Query Match:	54.80%	Indels:	0
DB:	22	Gaps:	0

US-09-770-509-2 (1-178) x AAH52021 (1-1140)

```

OY 4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 94 CTCMAAGGGGTGGAAATTCATCGGATCAACACCGCCCGGCGTTTGATGACCGAT 153
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 154 GCCGAGCTCAAACTGACGACGCGCGGCGGATCCACCGCGGGGCTGGCGCGCGCGAT 213
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 214 CCGGAGGTGCGCGGTAGGCGCGCGGAGGACCGCAAGACGAGATCGAAGAGCTGCTGCGC 273
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyLysThrCysThrGlyAlaAla 83
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 274 GGTGCGCAATGTTGTGTTCACCGCGCGGAGGGGGGCGGACCGGACCGGGGGGCA 333
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValAlaSerThr 103
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 334 CCGGTCGTCGCCACATCGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 393
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysLeu 123
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 394 CCGTTCCTGTCGAGGCGCAAGCGACGCAATCAGCCGCAAAATGGCATCGCGCGCGTG 453
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 124 AlaLysTyValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaAsp 143
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 454 CCGGAGAGTTGGGACACCCCTCATCGTATCCCAACGACCGGTTGCTCAAGTGGAGAT 513
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 144 LysSerThrThMetLeuGlnAlaPheArgTyValAlaAspValLeuLeuGluGlyVal 163
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 514 GCCGCGGTATCGGTGATGATGCTTCCGTAGCGCCGACGAGGTGCTCAACGCGGTG 573
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 574 CAGGGCATCACCGACTGATACCAACCGCGGCTCTAATCAACGTC 618
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 13
 AA199683/C
 ID AA199683 standard; DNA; 4403765 BP.

AC AA199683;
 XX
 DT 15-JAN-2002 (first entry)
 XX
 DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
 XX
 KW Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
 XX variation; epidemiology; patient treatment; epidemic monitoring; ds.
 OS Mycobacterium tuberculosis.
 XX
 FN US6294328-B1.
 XX
 PD 25-SEP-2001.
 XX
 PF 24-JUN-1998; 98US-0103840.
 XX
 PR 24-JUN-1998; 98US-0103840.
 XX

PA (GENO-) INST GENOMIC RES.
 XX
 PI Fleischmann RD, White OR, Fraser CM, Venter JC;
 XX
 DR WPI: 2001-647261/74.
 XX
 PT Evaluating strain variation of Mycobacterium tuberculosis, comprises
 PT determining the nucleotide sequence of the strain at positions in the
 PT genome corresponding to positions where M. tuberculosis strains CDC
 PT 1551 and H37Rv differ -
 XX
 XX
 PS Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
 XX
 CC The invention relates to evaluating strain variation within and between
 CC different populations of the tuberculosis bacterial pathogen,
 CC Mycobacterium tuberculosis or related Mycobacterium by determining the
 CC nucleotide sequence of the first strain at positions in the complete
 CC sequence of the genome that correspond to positions that differ in the
 CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AA199683) and
 CC H37Rv (AA199682). The method is useful for evaluating strain variation of
 CC M. tuberculosis and has valuable application in the fields of
 CC tuberculosis genetics, epidemiology, patient treatment and epidemic
 CC monitoring.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html?docID=6294328B1.
 XX
 SQ Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;

Alignment Scores:

Pred. No.:	3.97e-39	Length:	4403765
Score:	485.00 <td>Matches:</td> <td>95 </td>	Matches:	95
Percent Similarity:	74.29% <td>Conservative:</td> <td>35 </td>	Conservative:	35
Best Local Similarity:	54.29% <td>Mismatches:</td> <td>45 </td>	Mismatches:	45
Query Match:	54.80% <td>Indels:</td> <td>0 </td>	Indels:	0
DB:	22	Gaps:	0

US-09-770-509-2 (1-178) x AA199683 (1-4403765)

```

OY 4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2408087 CTCMAAGGGGTGGAAATTCATCGGATCAACACCGCCCGGCGTTTGATGACCGAT 2408028
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2408027 GCCGAGCTCAAACTGACGACGCGCGGATCCACCGCGGGGCTGGCGCGCGCGAT 2407968
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2407967 CCGGAGGTGCGCGGTAGGCGCGCGGAGGACGCAAGATCGAAGAGCTGCTGCGC 2407908
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyLysThrCysThrGlyAlaAla 83
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2407907 GGTGCGCAATGTTGTGTTCACCGCGCGGAGGGGGGCGGACCGGACCGGGGGGCA 2407848
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValAlaSerThr 103
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2407847 CCGGTGTCGTCGCCACATCGCCCGGACGCTGGGCGGCTGACCGGTGTCGACCCGG 2407788
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2407787 CCGTTCCTGTCGAGGCGCAACGCGACGATACGCGGCAAAATGGCATCGCGGCGTG 2407728
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 124 AlaLysTyValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaAsp 143
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2407727 CCGGAGAGTTCGACCTCATGATTCACCAACGACCGGTTCTTGAGATGGAGAT 2407668
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 144 LysSerThrThMetLeuGlnAlaPheArgTyValAlaAspValLeuLeuGluGlyVal 163
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2407667 GCCGCGGTATCGGTGATGATGCTTCCGTAGCGCGGACGAGGTGCTCAACGCGGTG 2407608
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

DB 2407607 CAGGGCATCACCACCTGATTACACCCCGGGTCTAATCAACGTC 2407563
RESULT 14
AA199682/C
ID AA199682 standard; DNA: 4411529 BP.
XX AA199682;
XX
XX 15-JAN-2002 (first entry)
XX
XX Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
DE
XX Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KM variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
XX Mycobacterium tuberculosis.
OS
XX US6294328-B1.
PN
XX 25-SEP-2001.
PD
XX 24-JUN-1998; 98US-0103840.
PF
XX 24-JUN-1998; 98US-0103840.
PR
XX (GENO-) INST GENOMIC RES.
XX
XX Fletschmann RD, White OR, Fraser CM, Venter JC;
PI
XX WPI: 2001-647261/74.
DR
XX
XX Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT determining the nucleotide sequence of the strain at positions in the
PT genome corresponding to positions where M. tuberculosis strains CDC
PT 1551 and H37Rv differ.
XX
XX Claim 3: SEQ ID NO 1; 3pp + Sequence Listing; English.
XX
XX The invention relates to evaluating strain variation within and between
CC different populations of the tuberculosis bacterial pathogen,
CC Mycobacterium tuberculosis or related Mycobacterium by determining the
CC nucleotide sequence of the first strain at positions in the complete
CC sequence of the genome that correspond to positions that differ in the
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AA199683) and
CC H37Rv (AA199682). The method is useful for evaluating strain variation of
CC M. tuberculosis and has valuable application in the fields of
CC tuberculosis genetics, epidemiology, patient treatment and epidemic
CC monitoring.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?docid=6294328B1.
CC
XX
SQ Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other:
Alignment Scores:
Pred. No.: 3, 98e-39 Length: 4411529
Score: 485.00 Matches: 95
Percent Similarity: 74.29% Conservative: 35
Best Local Similarity: 54.29% Mismatches: 45
Query Match: 54.80% Gaps: 0
DB: 22
US-09-770-509-2 (1-178) x AA199682 (1-4411529)
OY 4 LeungGLYValGluPheIleValAlaAsnThrAspGlnAlaLeuGLYArgSerLeu 23
DB 2409429 CTCGAAGGGGTGAATTCATCGGATGACACGACCCGAGGCTGTGATGAGCAT 2409370
OY 24 AlaProHlaLysIleThrLeuGLYValAspIleThrLysGluLeuGLYAlaGlySerLys 43
DB 2409369 GCGGACGTCAAACTGACGCGGCGGACACTCACCGGCGGCGGCGGCGGCGCAT 2409310
OY 44 ProGluLeuGLYValArgSerAlaGluGlnLysValAspIleGlnArgMetLeuGln 63

DB 2409309 CCGGAGGTGGCGGTAAAGCGCCGAGACGCCAAGACGACGANTCGAAGACTCTCGCC 2409250
OY 64 AspSerAsnMetLeuPheIleThrGlyMetGlyGlyCysThrGlyAlaAla 83
DB 2409249 GGTGCGCATGGTGTGTTGACACCGCGGAGGCGGAGAACCGCGGCGGCGCA 2409190
OY 84 ProValValAlaSerValAlaArgGluLeuGLYIleLeuThrValGlyValSerThr 103
DB 2409189 CCGGTGTCGCCACGATCGCCCGCAAGCTGGGCGCTGACGCTGGTGTGTCACCGC 2409130
OY 104 ProPheArgSerGluLeuProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
DB 2409129 CCGTCTCGTTGAGGAGCAAGCAGCAGCAATCAGCGCAAAATGCGATCGCGCGCTG 2409070
OY 124 AlaLysTyrValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaAsp 143
DB 2409069 CCGGAGAGTGGCACACCTCATCGATGATCCACACCGGTTGCTGACAGATGGAGAT 2409010
OY 144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspValLeuLeuGLYVal 163
DB 2409009 GCGCGCGTATCGCTGATGATGCTTCCGTAGCGCGCAGAGTGTCTCAACGCGCTG 2408950
OY 164 LysGlyValThrAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
DB 2408949 CAGGCGATCACCAGCCTGATTACCACCCGCGGTCTAATCAACGTC 2408905
RESULT 15
AAS52133
ID AAS52133 standard; DNA: 1172 BP.
XX
XX AAS52133;
AC
XX 13-FEB-2002 (first entry)
XX
XX Staphylococcus aureus DNA for cellular proliferation protein #550.
DE
XX Antisense; ds; prokaryotic cellular proliferation gene;
KM antibiotic; antibacterial; drug design.
KW
OS Staphylococcus aureus.
XX
XX WO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
XX
XX 21-MAR-2000; 2000US-191078P.
XX
XX 23-MAY-2000; 2000US-206848P.
XX
XX 26-MAY-2000; 2000US-207727P.
XX
XX 23-OCT-2000; 2000US-242578P.
XX
XX 27-NOV-2000; 2000US-253625P.
XX
XX 22-DEC-2000; 2000US-257931P.
XX
XX 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
PI
XX WPI: 2001-611495/70.
DR P-PSDB; AAO34274.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
PT
XX Claim 27: Seq ID No 4715; 511pp; English.
PS
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are

CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence encodes an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC

xx Sequence 1172 BP; 390 A; 194 C; 260 G; 328 T; 0 other;

Alignment Scores:

Pred. No.:	2,95e-43	Length:	1172
Score:	480.00	Matches:	95
Percent Similarity:	73.14%	Conservative:	33
Best Local Similarity:	54.29%	Mismatches:	47
Query Match:	54.24%	Indels:	0
DB:	23	Gaps:	0

US-09-770-509-2 (1-178) x AAS52133 (1-1172)

```

OY      4  LeuGIuGIyValGIuPhelIeValAlaSerThrAspCysGlnAlaLeuGIyArGSerLeu 23
DB      103 ATGAATTAATGTTGAATTTATCGCTATCAACACAGCGCTCAAGCTTAACCTTAATCTAAA 162
OY      24  AlaProHisLysIleThrLeuGIyLysAspIleThrLysGIyLeuGIyAlaGIySerLys 43
DB      163 GCTGAATCTAAATCCAAATCGGTGAATAAATTAAACAGTGTAGAGCAGAGACTAAT 222
OY      44  ProGIuLeuGIyLysArGSerAlaGIuGlnLysValAspIleGlnArgMetLeuGln 63
DB      223 CCTGAATCGTAAAAAAGCTGCAGAGAACTCTGTGAACAAATTTGAATGCAATCCAA 282
OY      64  AspSerAsnMetLeuPhelIeThrGIyMetGIyLysGIyThrCysThrGIyAlaAla 83
DB      283 GGTGCAGACATGATTTGTTACTTCTGATGAGGTGGCGAAGCTGTAAGTGTGACGA 342
OY      84  ProValValAlaSerValAlaArGLeuGIyIleLeuThrValGIyValValSerThr 103
DB      343 CCAGTCGTGCTAAATTCGAAGAAGAAATGGCGCAATTAAGTGTGTGTAAGTCTGT 402
OY      104  PropheArgSerGIyLysProAsnArgThrArgLeuAlaAsnAlaGIyValLysGIuLeu 123
DB      403 CCATTAGTTTGAAGAGCTAAACGTCACAACTCAACCTGCTGCTGAGTAGAAGCTATG 462
OY      124  AlaLysTyValAspThrLeuIleValProAsnGlnAsnLeuAlaLeuAlaAsp 143
DB      463 AAAGCTGCAGTATGATTAATTCGTTATACCAATGACCGTTTATTAAGATATCGTTGAC 522
OY      144  LysSerThrThrMetLeuGIuAlaPheArgTyValAspAspValLeuLeuGIuGIyVal 163
DB      523 AAATCTACCGCAATGATGAAGACATTAAGAAGAGCTGACAAACGTGTACGCCAAGGTGTA 582
OY      164  LysGIyValThrAspLeuIleValArgProGIyLeuIleAsnLeu 178
DB      583 CAAGGTATCTCAGACTTAATCGCTGTTCTGTAAGTAAGTAACCTTA 627

```

Search completed: June 2, 2003, 14:34:50
 Job time : 20511 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 2, 2003, 08:48:49 ; Search time 57 Seconds
(Without alignments)
957.692 Million cell updates/sec

Title: US-09-770-509-2

Perfect score: 885
Sequence: 1 ASOLEGVEFIVANTDCALG.....LLEGVKGVTDLIVRPLINL 178

Scoring table: BLOSUM62
Xgapop 10.0 , Ygapext 0.5
Ygapop 6.0 , Ygapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame-p2n.model -DEV=x1h
-O=/cg22.1/USPTO.spool/US09770509/runat_27052003_083650_18840/app.query.fasta_1.327
-DB=Issued_Patents_NA -OPMT=fastap -SOFTX=tri -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=300 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09770509 @CGN1.1.40 @runat_27052003_083650_18840 -NCPU=6 -ICPU=3
-NO_XMAP -NO_XMAP -LARGEQUERY -NEG_SCORES=0 -MAT -DSPLOCK=100 -LONLOC
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=6 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued_Patents_NA.*

- 1: /cg22.6/ptodata/1/1na/5A.COMB.seq:*
- 2: /cg22.6/ptodata/1/1na/5B.COMB.seq:*
- 3: /cg22.6/ptodata/1/1na/6A.COMB.seq:*
- 4: /cg22.6/ptodata/1/1na/6B.COMB.seq:*
- 5: /cg22.6/ptodata/1/1na/6C.COMB.seq:*
- 6: /cg22.6/ptodata/1/1na/6D.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	485	54.8	4403765	4 US-09-103-840A-2	Sequence 2, Appl1
C 2	485	54.8	4411529	4 US-09-103-840A-1	Sequence 1, Appl1
C 3	480	54.2	1185	4 US-09-134-001C-1997	Sequence 1997, Ap
C 4	472	53.3	1005	4 US-08-987-146-3	Sequence 3, Appl1
C 5	472	53.3	1008	4 US-08-987-146-1	Sequence 1, Appl1
C 6	472	53.3	1260	4 US-09-120-426-1	Sequence 4, Appl1
C 7	472	53.3	2702	4 US-08-987-146-4	Sequence 1, Appl1
C 8	472	53.3	21706	4 US-08-961-527-36	Sequence 36, Appl1
C 9	464	52.4	7143	4 US-09-381-862-4	Sequence 4, Appl1
C 10	463	52.3	1638	2 US-08-883-515-3	Sequence 3, Appl1
C 11	429	48.5	1425	2 US-08-883-515-1	Sequence 3, Appl1
C 12	376	42.5	396	4 US-09-134-001C-1996	Sequence 1996, Ap

13	369.5	41.8	665	4 US-09-221-017B-658	Sequence 658, App
14	271	30.6	850	4 US-09-120-426-3	Sequence 3, Appl1
15	196	22.1	859	4 US-08-961-083-123	Sequence 123, App
16	90.5	10.2	7100	4 US-09-308-375-1	Sequence 1, Appl1
C 17	85.5	9.7	2245	4 US-08-617-801A-3	Sequence 3, Appl1
C 18	83	9.4	1888	1 US-08-229-145-13	Sequence 13, Appl1
C 19	81.5	9.2	4403765	4 US-09-103-840A-2	Sequence 2, Appl1
C 20	81.5	9.2	4411529	4 US-09-103-840A-1	Sequence 1, Appl1
C 21	80.5	9.1	1621	3 US-09-013-881-14	Sequence 14, Appl1
C 22	79	8.9	1167	4 US-09-134-001C-2114	Sequence 2114, Ap
C 23	78.5	8.9	1574	4 US-09-189-527-10	Sequence 10, Appl1
C 24	78.5	8.9	9244	4 US-08-961-527-68	Sequence 68, Appl1
C 25	78	8.8	2109	2 US-08-617-801A-5	Sequence 5, Appl1
C 26	78	8.8	8906	2 US-08-826-267-1	Sequence 1, Appl1
C 27	77.5	8.8	1946	2 US-08-861-464-1	Sequence 1, Appl1
C 28	77.5	8.8	1946	2 US-08-396-001-1	Sequence 1, Appl1
C 29	77.5	8.8	1946	4 US-09-323-433A-1	Sequence 1, Appl1
C 30	76.5	8.6	1269	4 US-09-134-001C-197	Sequence 1, Appl1
C 31	76.5	8.6	2844	4 US-09-257-703-2	Sequence 2, Appl1
C 32	76.5	8.6	3156	2 US-08-887-518-1	Sequence 1, Appl1
C 33	76.5	8.6	3156	2 US-08-887-518-1	Sequence 1, Appl1
C 34	76.5	8.6	3156	2 US-09-032-475-1	Sequence 1, Appl1
C 35	75.5	8.5	1229	1 US-08-440-861-1	Sequence 1, Appl1
C 36	75.5	8.5	1229	1 US-08-433-854-1	Sequence 1, Appl1
C 37	75.5	8.5	1229	1 US-08-174-745A-1	Sequence 1, Appl1
C 38	75.5	8.5	1229	2 US-08-195-947-1	Sequence 1, Appl1
C 39	75.5	8.5	1229	2 US-08-433-885-1	Sequence 1, Appl1
C 40	75.5	8.5	1229	2 US-08-433-908B-1	Sequence 1, Appl1
C 41	75.5	8.5	1229	4 US-08-410-614-1	Sequence 1, Appl1
C 42	75	8.5	2504	1 US-08-073-384C-3	Sequence 3, Appl1
C 43	75	8.5	2504	1 US-08-254-359A-3	Sequence 3, Appl1
C 44	75	8.5	2504	1 US-08-483-043-3	Sequence 3, Appl1
C 45	75	8.5	2504	1 US-08-481-238-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VEMMER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
Alignment Scores:
Pred. NO.: 1.02e-46
Score: 485.00
Percent Similarity: 74.29%
Best Local Similarity: 54.29%
Query Match: 54.80%
DB: 4
Gaps: 0
US-09-770-509-2 (1-178) x US-09-103-840A-2 (1-4403765)

OY	4	LeuGIuGIValIaIaIuPheIIeValaIaIaAspThrAspCysGlnaIaIeucIaIaArgSerLeu	23
Db	2408087	CTCAAAAGCGGTGGAAATTCATCGCGATCAACACACCGCCGAGCGGTGTGGATGAAGCGAT	24080288
OY	24	AlaProHizLysLileThrLeuGIuLysAspIleThrLysGIuLeucIValaIaGIuSerLys	43
Db	2408027	GGCGAGCTGAACCTGCAGCTGGCGCGGAGCTCCACCCCGGGCTGGGCGCGCGCGGAT	24079688
OY	44	ProGIuLeuGIuLysArgSerAlaGIuGlnGlnLysValaAspIleGlnArgMetLeuGln	63
Db	2407967	CCGGAGGTGGCGCTGAAGGCCCGCCGAGAGCCCGAAGACAGAAATGCAACACTGCTGGCG	24079088
OY	64	AspSerAsnMetLeuPheIleThrGIuLysMetGIuGIuLysThrCysThrGIuValaIa	83
Db	2407907	GGTGGCGAATGATGTTGTTCACCGCGGCGAGGGGGGCGGAACCGGCGACCGGGGGGCA	24078488
OY	84	ProValValaIaSerValaIaArgGlnGlnLysIleLeuThrValaGIuValaIaSerThr	103
Db	2407847	CCCGTGTGCGCAAGCATCGCCCGGAAAGCTGGCGCGTGTGACCTGGTGTGGTCACCGG	24077888
OY	104	ProPheArgSerGIuGIuProAsnArgThrArgLeuIaAsnIaGIuValaLysGIuLeu	123
Db	2407787	CCGTTCTCGTTCAGAGGGCAAGCGACGACAAATAGGCCGAAATAGGCATCGCGCGCTG	24077288
OY	124	AlaLysThrValaAspIleuIleValaIaProAsnGlnAsnLeuLeuIaIaLeuAlaAsp	143
Db	2407727	CGGAGAGCTTGCACACCCCTCATCTGTATATCCACAGCACCGGCTGTGCAGATGGAGAT	24076688
OY	144	LysSerThrThrMetLeuGlnaIaIaPheArgThrIaAspAspValaIeucGIuGIuVala	163
Db	2407667	GGCGGGGATCGCTGATGATGATGCTTCCGTAGCGCCGACAGAGGTGCTGCTCAACGGCGTG	24076088
OY	164	LysGIuValaIaThrAspLeuIleValaArgProGIuLeuIleAsnLeu	178
Db	2407607	CAGGGATACCGACACTGATTACACACCCCGGGTCAATCAACACTG	2407563

```

RESULT 2
US-09-103-840A-1/c
: Sequence 1, Application US/09103840A
: Patent No. 62364328
:
GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: FILE REFERENCE: 24366-20007.00
: CURRENT APPLICATION NUMBER: US/09/103,840A
: CURRENT FILING DATE: 1998-06-24
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: Patentin Ver. 2.1
:
SEQ ID NO 1
:
: LENGTH: 4411529
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.:          1.02e-46          Length:          4411529
Score:              485.00             Matches:         95
Percent Similarity: 74.29%             Conservative:    35
Best Local Similarity: 54.29%           Mismatches:     45
Query Match:        54.80%             Indels:         0
DB:
      4
      gaps:      0
US-09-770-509-2 (1-178) x US-09-103-840A-1 (1-4411529)

```

QY	4	leuglucilvAlGupheileAlaAsnrrAspcysginalenuglyArgserLeu	23
Db	2409429	CTCAAGGCGTGAATTCATCGCGATCAACACCGAGCCCGCGGTGTGATGACGAT	2409370

OY	24	AlaProHIsLysLLeThrLeuGLysAspLLeuThrLysGLLeuGLyAlaLysSerLys	43
Db	2409369	GGCGAGCTCAAACTCGACGTCGGCGGCGACTCCACCCCGGGCTGGGGCCGCCGAT	2409310
OY	44	ProGluLeuGLyLysArgSerAlaGluGlnLysValAspLLeuInArgMetLeuGln	63
Db	2409309	CCGAGAGTCGGCCGTAGGGCCGCGAGGACCCAGAGCATGAGTAGCTGGCGC	2409250
OY	64	AspSerAsmMetLeuPheLLeThrGlyGlyMetGlyGlyLysThrCysThrGlyAlaAla	83
Db	2409249	GGTGGCGACATGGTGTGTTCACCGCCGCGCAAGGGGGCGGAAACCGCCGCGGGGGGCA	2409190
OY	84	ProValValAlaSerValAlaArgGluLeuGlyLLeuLeuThValGlyValValSerThr	103
Db	2409189	CCCGTCGTGCCACGATCGGCCGCGAGCTGGGGCGGTGGACCGTGGTGTGGTCACCGG	2409130
OY	104	ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu	123
Db	2409129	CCGTCTCTGCTTCAGGGCGAAGCGACGACATCGAGCGGCGGAAATGGCTCGGGCGCTG	2409070
OY	124	AlaLysTyValAspThrLeuLLeValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp	143
Db	2409069	CGGGAGAGTTGGCAGACCCCTCATCTGATTTCCCAACGACCGGTTCGTGCAGATGGGAGAT	2409010
OY	144	LysSerThrThrMetLeuGlnAlaAlaPheArgTyAlaAspAspValLeuLeuGluGlyVal	163
Db	2409009	GGCGCGATCGCTGATGATGATGCTTCCGTAGCGCGGACGAGAGTCTGCTCAACGGCGTG	2408950
OY	164	LysGlyValThrAspLeuLLeValArgProGlyLeuLLeAsnLeu	178
Db	2408949	CAGGGCATACCGACCTGATTTACCAACCCCGGGGTCTTAATCAACGTC	2408905

```

1 RESULT 3
2 US-09-134-001C-1997
3 Sequence 1997 Application US/09134001C
4 Patent No. 6380370
5 GENERAL INFORMATION:
6 APPLICANT: Lynn Doucette-Stamm et al
7 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
8 TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
9 FILE REFERENCE: GTC-007
10 CURRENT APPLICATION NUMBER: US/09/134,001C
11 CURRENT FILING DATE: 1998-08-13
12 PRIOR APPLICATION NUMBER: US 60/064,964
13 PRIOR FILING DATE: 1997-11-08
14 PRIOR APPLICATION NUMBER: US 60/055,779
15 PRIOR FILING DATE: 1997-08-14
16 NUMBER OF SEQ ID NOS: 5674
17 SEQ ID NO 1997
18 LENGTH: 1185
19 TYPE: DNA
20 ORGANISM: Staphylococcus epidermidis
21 US-09-134-001C-1997

```

Alignment scores:	
Pred. No.:	2.32e-51
Score:	480.00
Percent Similarity:	73.14%
Best Local Similarity:	54.29%
Query Match:	54.24%
DB:	4
Gaps:	0

US-09-770-509-2 (1-178) x US-09-134-001C-1997 (1-1185)

Oy 4 LeucGluGlyValGlnPheIleValAlaasnThrAspCysGlnAlaLeuGlyArgSerLeu 23
 :: | | | | | | | | | | | | | |
Db 103 ATGATAATATGTGAATTATTTCACATCAACACTGAGTGGACAACCTTTAAACTTATCAAAA 167

OY 24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerIys 43
||| ||||| ::||::: :||||:|||||||:::
Db 163 GCTGAATCAAAATTTCAAATTGGTGTAGAAATTAAACGTGATTTAGTGCGTGCCTAAT 22

OY 44 ProGluLeuGlyLysArgSerAlaGluGlnGlnIleValAspIleGlnArgMetLeuGln 63

FILING DATE: December 8, 1997
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Webster, Thomas D.
 REGISTRATION NUMBER: 39,872
 REFERENCE/DOCKET NUMBER: X-11755
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 317-276-3334
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1008 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOHETICAL: NO
 ANTI-SENSE: NO
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1005
 US-08-987-146-1

Alignment Scores:
 Pred. No.: 1 88e-50 Length: 1008
 Score: 472.00 Matches: 91
 Percent Similarity: 72.83% Conservative: 35
 Best Local Similarity: 52.60% Mismatches: 47
 Query Match: 53.33% Indels: 0
 DB: 4 Gaps: 0

US-09-770-509-2 (1-178) x US-08-987-146-1 (1-1008)

QY 6 GlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaPro 25
 DB 112 GCGGTAAATTATTCGACGAACACAGATGTACAGCATGTAGTACAAACAACTGAG 171
 QY 26 HisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGlu 45
 DB 172 ACTGTATTACAGTTCAGTGGACCTAAATGACCTGCTGGTGGCGAGGAGGCAACCTGAG 231
 QY 46 LeuGlyLysArgSerAlaGluGlnGlnIleValAspIleGlnArgMetLeuGlnAspSer 65
 DB 232 GTTGGCTGTAAGCCCGCTGAAGAAAGCAAAACATGACGGAAGCTATTAGTGTGCGC 291
 QY 66 AsnMetLeuPheIleThrGlyMetGlyGlyLysThrCysThrGlyAlaAlaProVal 85
 DB 232 GATATGCTTCATCATCAGTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 351
 QY 86 ValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPhe 105
 DB 352 ATTTGCTCGTATCGCAAGATTGAGTGGCTTACAGTTGCTGTGTAACAGCTCCCTTT 411
 QY 106 ArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLys 125
 DB 412 GGTTTGAAGGAAGTAAAGCGTGACAAATTGCTGTACAAAGAAATCAATCACTTCGTGAG 471
 QY 126 TyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSer 145
 DB 472 CATGTAGACACTCTATTGATTATCTCAACAAATTTGCTTGAATTGTGATTAAGAA 531
 QY 146 ThrThrMetLeuGlnAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGly 165
 DB 532 ACACCGCTTTTGGAGGCTCTTACGGAAGCGGATTAACGTTCTTCGCAAGGTTCAGAGG 591
 QY 166 ValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
 DB 592 ATTACGATTGATTACCAATCCAGATTTGATTAACTT 630

RESULT 6
 US-09-120-426-1
 ; Sequence 1, Application us/09120426
 ; Patent No. 6197300
 ; GENERAL INFORMATION:

APPLICANT: Fueyo, Joanna Lynn
 APPLICANT: Lonetto, Michael A.
 APPLICANT: Pearce, Kenneth
 TITLE OR INVENTION: ftsZ
 FILE REFERENCE: GM10068
 CURRENT APPLICATION NUMBER: US/09/120,426
 CURRENT FILING DATE: 1998-07-22
 EARLIER APPLICATION NUMBER: 60/055,720
 EARLIER FILING DATE: 1997-08-12
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 1
 LENGTH: 1260
 TYPE: DNA
 ORGANISM: Streptococcus pneumoniae
 FEATURE:

US-09-120-426-1

Alignment Scores:
 Pred. No.: 2.61e-50 Length: 1260
 Score: 472.00 Matches: 91
 Percent Similarity: 72.83% Conservative: 35
 Best Local Similarity: 52.60% Mismatches: 47
 Query Match: 53.33% Indels: 0
 DB: 4 Gaps: 0

US-09-770-509-2 (1-178) x US-09-120-426-1 (1-1260)

QY 6 GlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaPro 25
 DB 112 GCGGTAAATTATTCGACGAACACAGATGTACAGCATGTAGTACAAACAAAGCTGAG 171
 QY 26 HisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGlu 45
 DB 172 ACTGTATTACAGTTCAGTGGACCTAAATGACCTGCTGGTGGCGAGGAGGAGGAGGAGGAGG 231
 QY 46 LeuGlyLysArgSerAlaGluGlnGlnIleValAspIleGlnArgMetLeuGlnAspSer 65
 DB 232 GTTGGCTGTAAGCCCGCTGAAGAAAGCAAAACATGACGGAAGCTATTAGTGTGCGC 291
 QY 66 AsnMetLeuPheIleThrGlyMetGlyGlyLysThrCysThrGlyAlaAlaProVal 85
 DB 292 GATATGCTTCATCATCAGTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 351
 QY 86 ValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPhe 105
 DB 352 ATTTGCTCGTATCGCAAGATTGAGTGGCTTACAGTTGCTGTGTAACAGCTCCCTTT 411
 QY 106 ArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLys 125
 DB 412 GGTTTGAAGGAAGTAAAGCGTGACAAATTGCTGTGTAACAGCTCCCTTT 471
 QY 126 TyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSer 145
 DB 472 CATGTAGACACTCTATTGATTATCTCAACAAATTTGCTTGAATTGTGATTAAGAA 531
 QY 146 ThrThrMetLeuGlnAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGly 165
 DB 532 ACACCGCTTTTGGAGGCTCTTACGGAAGCGGATTAACGTTCTTCGCAAGGTTCAGAGG 591
 QY 166 ValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
 DB 592 ATTACGATTGATTACCAATCCAGATTTGATTAACTT 630

RESULT 7
 US-08-987-146-4
 ; Sequence 4, Application US/08987146
 ; Patent No. 6350866
 ; GENERAL INFORMATION:
 APPLICANT: Skatrud, Paul L.
 APPLICANT: Peery, Robert B.
 APPLICANT: Rocky, Pamela K.
 APPLICANT: Wang, Q. May

APPLICANT: Rostack Jr., Paul R.
TITLE OF INVENTION: Streptococcus Pneumoniae Gene Sequence
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/987,146
FILING DATE: December 8, 1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-11755
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2702 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-987-146-4

Alignment Scores:
Pred. No.: 8,07e-50 Length: 2702
Score: 472.00 Matches: 91
Percent Similarity: 72.83% Conservative: 35
Best Local Similarity: 52.60% Mismatches: 47
Query Match: 53.33% Indels: 0
Gaps: 0

US-09-770-509-2 (1-178) x US-08-987-146-4 (1-2702)

QY 6 GtYvAlGluPhelIeVAlAlaSnThrAsPcYsGlnAlaLeuGlyArGserLeuAlaPro 25
Db 170 GCGCTGAATTTATCCAGCAACACAGATGTACCAACATTGACTACAAAAGCTGAG 229
QY 26 HtSLyLleThrLeuGlyLysAspIleThrLysGlyLeuGlyLysSerLeuGlu 45
Db 230 ACTGTTATTCAGTTGGAGCACTTAATGACTGCTGGTTGGTGAGAGGCTCAACCTGAG 289
QY 46 LeuGlyLysArGserAlaGluGlnGlyValAspIleGlnArGmetLeuGlnAspSer 65
Db 290 GTTGGTCGTAAGCCGCTGAAGAGAAAGCAAACTGACGGAAGCTATTAGGGGCC 349
QY 66 AaMleuPheIleThrLysGlyMetGlyGlyThyrGlyStrGlyAlaAlaProVal 85
Db 350 GATATGCTCTTACTGCTGCTGATGGAGGAGGCTGGAAGCTGAGCTGCTGCTT 409
QY 86 ValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValIleSerThrProPhe 105
Db 410 ATTGCTCGATGCCCAAGATTAGGTGGCTTACCTTGTGTTGTAACGCTCCCTT 469
QY 106 ArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLys 125
Db 470 GGTTTGAAGAGTAAGCGTGAACAATTGCTGTAGACGAATCAATCACTTCGAG 529
QY 126 TyrValAspThrLeuIleValAlaProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSer 145

Db 530 CATGTAGACACTCTATTGATTATCTCAACAAACATTTGCTGAATTTGTATAGAAA 589
QY 146 ThrThrMetLeuGlnAlaPheArgTyrAlaAspValLeuLeuGlyValLysGly 165
Db 590 ACACCGCTTTGGAGGCTCTTACCGAAGCGGATACGTTGCTCAAGGTTCACAGG 649
QY 166 ValThrasPleuIleValArGProGlyLeuIleAsnLeu 178
Db 650 ATTACGATTGTATTACCAATCCAGATTGATTACCTT 688

RESULT 8
US-08-961-527-36
Sequence 36, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 21706 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-36

Alignment Scores:
Pred. No.: 1.75e-48 Length: 21706
Score: 472.00 Matches: 91
Percent Similarity: 72.83% Conservative: 35
Best Local Similarity: 52.60% Mismatches: 47
Query Match: 53.33% Indels: 0
Gaps: 0

US-09-770-509-2 (1-178) x US-08-961-527-36 (1-21706)

QY 6 GtYvAlGluPhelIeVAlAlaSnThrAsPcYsGlnAlaLeuGlyArGserLeuAlaPro 25
Db 2761 GCGCTGAATTTATCCAGCAACACAGATGTACCAACATTGACTACAAAAGCTGAG 2820
QY 26 HtSLyLleThrLeuGlyLysAspIleThrLysGlyLeuGlyLysSerLeuGlu 45
Db 2821 ACTGTTATTCAGTTGGAGCACTTAATGACTGCTGGTTGGTGAGAGGCTCAACCTGAG 2880
QY 46 LeuGlyLysArGserAlaGluGlnGlyValAspIleGlnArGmetLeuGlnAspSer 65
Db 2881 GTTGGTCGTAAGCCGCTGAAGAGAAAGCAAACTGACGGAAGCTATTAGTGTGCTC 2940

```

Oy      66 AsnMetLeuPheIleThrGlyMetGlyGlyThrCysThrGlyAlaAlaProVal 85
       : :::::::::::::::::::: |::::::::::::::::::| ::::::::::::::
Db      2941 GATATGCTCTTCACTACTGCTGTGATGGAGAGGCCTCTGGAAC TCGAAGCTGCTCTT 3000
       : ::::::::::::::::::::|::::::::::::::::::| ::::::::::::::
Oy      86 ValAlaSerValAlaArgJugLeuGlyLlleuThrValGlyValValSerThrProPhe 105
       : ::::| ::::::::::::::| ::::::::::::::| ::::::::::::::| ::::
Db      3001 AATGCTCGTATCCCAAAGATTATTTAGTGCGCTTACAGATGTGTTGTAACAACGCCCTT 3060
       : ::::::::::::::::::::| ::::::::::::::::::::| ::::::::::::::
Oy      106 ArgSerGluGlyProAsnArgThrArgLeuGlnAlaSnAlaGlyValLlysGluLeuAlaLys 125
       : ||||| ::::| ::::| ::::| ::::| ::::| ::::| ::::| ::::| ::::
Db      3061 GCTTTTGAAGGAGTAAAGCGTCGACAAATTTCTCTAGAGAAGATCATCAATCACTTCGTAG 3120
       : ::::::::::::::::::::| ::::::::::::::::::::| ::::::::::::::
Oy      126 TyrValAspThrLeuIleValAlaProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSer 145
       : ::::::::::::::::::::| ::::::::::::::::::::| ::::::::::::::
Db      3121 CATGTAGACACCTCTATGATTATTCATAACAACAATTTCTCTGAAATTTGTGTGATTAAGAAA 3180
       : ::::::::::::::::::::| ::::::::::::::::::::| ::::::::::::::
Oy      146 ThrThrMetLeuGluAlaPheArgTYrAlaAspAspValLeuLeuGluGlyValLysGly 165
       : ::::| ::::::::::::::| ::::::::::::::| ::::::::::::::| ::::
Db      3181 ACACCGCTTTGGAGGCTCTTAGCGAAGCGGATAACGTTCTTCGNCACAGGTGTCCAAGG 3240
       : ::::::::::::::::::::| ::::::::::::::::::::| ::::::::::::::
Oy      166 ValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
       : ::::::::::::::::::::| ::::::::::::::::::::| ::::::::::::::
Db      3241 ATTACCGATTGATTACCAAATCCAGATTGATTAACTT 3279
       : ::::::::::::::::::::| ::::::::::::::::::::| ::::::::::::::

RESULT 9
US-09-381-862-4
; Sequence 4, Application US/09381862
; Patent No. 6245906

GENERAL INFORMATION:
APPLICANT: Ueyama, Hiroshi
APPLICANT: Abe, Kanako
APPLICANT: Keshi, Hiroyuki
APPLICANT: Matsuhisa, Akio
TITLE OF INVENTION: PROBES FOR THE DIAGNOSIS OF INFECTIONS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/381,862
FILING DATE:
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 1997-71077
FILING DATE: 25-MAR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/Jp98/01288
FILING DATE: 23-MAR-1998
ATTORNEY/AGENT INFORMATION:
NAME: Cawley, Jr., Thomas A.
REGISTRATION NUMBER: 40,944
REFERENCE/DOCKET NUMBER: 19036/36274
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 7143 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

```

```

; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
; STRAIN: Clinical Isolate SP-26-36
US-09-381-862-4

Alignment Scores:
Pred. NO.: 3.49e-48 Length: 7143
Score: 464.00 Matches: 93
Percent Similarity: 71.10% Conservative: 30
Best Local Similarity: 53.76% Mismatches: 50
Query Match: 52.43% Indels: 0
DB: 4 Gaps: 0

US-09-770-509-2 (1-178) x US-09-381-862-4 (1-7143)

OY 6 GLYValGluPheIleValAlaSnThrAspCysGlnAlaLeuGlyArgSerLeuAlaPro 25
    |||||
DB 3673 GGTGTGAGTTTCATCGCGCAAAATACAGACTTACAGCATTAAGCTATCAACAAAGCTGAA 3732
    |||||

OY 26 HisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGlu 45
    |||||
DB 3733 ACGGTTATTCAACTAGAGCCCTAAATTAATCACTGCTGAGACTGTGCTGGAGCAACCTGAA 3792
    |||||

OY 46 LeuGlyAsnArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSer 65
    |||||
DB 3793 GTAGAGACTAAAGCTGCTGGAAGAAAGCAAAATTTACACAGAACTCTTACAGAGCG 3852
    |||||

OY 66 AsnMetLeuPheIleThrGlyGlyMetGlyGlyLysThrCysThrGlyAlaAlaProVal 85
    |||||
DB 3853 GATATGGATTATTATTAAGTCCGCGATGGTGCTGCTGGAGACAGGCGCTGCACCGGTT 3912
    |||||

OY 86 ValAlaSerValAlaIaArgGluLeuGlyLysIleLeuThrValGlyValSerThrProPhe 105
    |||||
DB 3913 ATTCTCGTACGCTAGAGAGTTGGGAGCCCTGACAGTAGCTGTGTTACTCCGCCGCTT 3972
    |||||

OY 106 ArgSerGluGlyProAsnArgThrArgLeuAlaSnAlaGlyValLysGluLeuAlaLys 125
    |||||
DB 3973 GGTITTGAGGTACCAAAACGCTGAATTTTCTATTGAGAGTATCGAAGAACTCCGTGAA 4032
    |||||

OY 126 TyrValAspThrLeuIleValAlaProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSer 145
    |||||
DB 4033 CAATGTGACTTGTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4092
    |||||

OY 146 ThrThrMetLeuGluAlaPheArgTyrValAlaAspValLeuLeuGluGlyValLysGly 165
    |||||
DB 4093 ACACCTTATTATAGAGCACTTAGTGAAGCGTATATGTTTACGTGACGAGGATTCAGAGG 4152
    |||||

OY 166 ValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
    |||||
DB 4153 ATATCCGACTTAATTACTAGTCCGCGGCTTATCATCTC 4191
    |||||

RESULT 10
US-08-883-515-3
; Sequence 3, Application US/08883515
; Patent NO. 5981836
; GENERAL INFORMATION:
; APPLICANT: Ostryoung, Katherine W
; TITLE OF INVENTION: PLANT CHLOROPLAST DIVISION GENES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Plokeney Street
; City: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/883,515

```

FILED DATE: 800
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 920905.90016
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1628 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..1316
US-08-883-515-3

Alignment Scores:
Pred. No.: 5 26e-49 Length: 1628
Score: 463.00 Matches: 89
Percent Similarity: 69.83% Conservative: 36
Best Local Similarity: 49.72% Mismatches: 52
Query Match: 52.32% Indels: 2
DB: 2 Gaps: 1

US-09-770-509-2 (1-178) x US-08-883-515-3 (1-1628)

QY 2 SerGlnLeuGluGlyValAlaGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
DB 297 ACTGAATGTCAGAGTGTGAGTGTCTGATGTCACACGATATATCCAGGCGATGAGATG 356
QY 22 Ser-----LeuAlaProHsLysIleThrLeuGlyLysAspIleThrLysGlyLeuGly 39
DB 357 TCTCCTGTTTGGCTGATAGTATAGTACAAATGTTGAGAGAGTGTGAGGCGTTAGCT 416
QY 40 AlaGlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnValAlaPheGln 59
DB 417 GCTGGAGGAATCCAGAAATCGTATGATGCTGCTGAGACAGCAAGAAAGATTATGAA 476
QY 60 ArgMetLeuGlnAspSerAsnMetLeuPheIleThrLysIleMetGlyGlyThrCys 79
DB 477 GAAGCCTTATGAGTCAGATATGCTGTTGTCACAGCTGGAATGGCGGTGAACTGGC 536
QY 80 ThrGlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGly 99
DB 537 ACTGGGACAGCCCTGTAATTCAGAGAAATGCAAGGCGATGATATGACAGTGGT 596
QY 100 ValValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGly 119
DB 597 ATGGCCACAAACCGCTTCTGCTTGAAGGCGCAAGAACTGTCAGCCTCAAGAAAGG 656
QY 120 ValLysGluLeuAlaLysTyrValAspThrLeuIleValAlaProAsnGlnAsnLeu 139
DB 657 CTTCGATCTCTCAGACACATGTCACACTCATCTCATTCCAATGACAACTGCTT 716
QY 140 AlaLeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeu 159
DB 717 ACAGCTGCTCTCAGTCACTCGGTTACAGAGCAATTTATCTAGCTGATGATATCTC 776
QY 160 LeuGluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
DB 777 GGTACGGGGGTTGCTGATATCTGATATCATTCATTCCTGTTGGTGAATGTC 833

RESULT 11
US-08-883-515-1
GENERAL INFORMATION:
APPLICANT: Austlin-Phillips, Sandra
APPLICANT: Burgess, Richard R.
APPLICANT: German, Thomas L.

APPLICANT: Ziegelhofer, Thomas L.
TITLE OF INVENTION: Transgenic Plants as an Alternative
SOURCE OF INVENTION: Source of Lignocellulosic-Degrading Enzymes
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Ste. 401
CITY: Madison
STATE: Wisconsin
COUNTRY: U.S.A.
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,495
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/028,718
FILING DATE: 17-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09820.036
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
Sequence 1, Application US/08883515
Patent No. 5981836

GENERAL INFORMATION:
APPLICANT: Osteryoung, Katherine W
TITLE OF INVENTION: PLANT CHLOROPLAST DIVISION GENES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,515

FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 920905.90016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1425 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 25..1326
US-08-883-515-1

Alignment Scores: 8.68e-45 Length: 1425
Pred. No.: 8.68e-45

Score: 429.00 Matches: 86
 Percent Similarity: 70.79% Conservative: 40
 Best Local Similarity: 48.31% Mismatches: 52
 Query Match: 48.47% Indels: 0
 DB: 2 Gaps: 0

US-09-770-509-2 (1-178) x US-08-883-515-1 (1-1425)

QY 1 A1aSerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGly 20
 DB 304 TCNAAGGGTTTACAGAGTGTGATTTCTAAGCGATTAACAGGATTCACAGCTCTGTTA 363
 QY 21 ArgSerLeuAlaProHisIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAla 40
 DB 364 CAGTTTCTGCTAGAACCCACTTCAATGAGAACTTTTAACTCGGCGCTTGACACT 423
 QY 41 GlySerLysProGluLeuGlyLysArgSerAlaGluGlnIleLysValAspIleGlnArg 60
 DB 424 GGTGGAAACCCGCTTCTTGAGAACACAGCTCCAGAACATCAAAAGATGCAATTGCTAAT 483
 QY 61 MetLeuGlnAspSerAsnMetLeuPheIleThrGlyMetGlyGlyLysThrCysThr 80
 DB 484 GCTCTTAAAGATGACAGCTTTGTTCACTGCTGATGGGTGGTGAACAGAGTCT 543
 QY 81 GluAlaAlaProValAlaAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyAl 100
 DB 544 GGTGCTGACCTGCTGTAGTACAGATTTGCAAGGATGCTGTTATTGACGTTGTGT 603
 QY 101 ValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
 DB 604 GTTACCTATCCGTTTACGTTTGAAGACGCAAAAGATCTTGGACGCGACTGGAAGCTATT 663
 QY 121 LysGluLeuAlaLysTyrValAspThrLeuIleValProAsnGlnAsnLeuAla 140
 DB 664 GAAAGCTCCAAAGATGTGATACCTTATGCTGATTCGCAAAAGATCGTCTAGAT 723
 QY 141 LeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspValLeuLeu 160
 DB 724 ATGTGCTAGAACACAGCGCACTTACAGCGCGTTTCTTCTTGCAGATGATGTTTACGC 783
 QY 161 GluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuLeuLeu 178
 DB 784 CAAAGAGTACAGAGATCTCAGATATTATTAATACTAGCTGACTGACATGATGTG 837

RESULT 12

US-09-134-001C-1996/C
 ; Sequence 1996, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134, 001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 1996
 ; LENGTH: 396
 ; TYPE: DNA
 ; ORGANISM: Staphylococcus epidermidis
 ; US-09-134-001C-1996

Alignment Scores:

Score: 6.68e-39 Length: 396
 Percent Similarity: 376.00 Matches: 72
 Best Local Similarity: 55.38 Conservative: 27
 Query Match: 42.49% Mismatches: 31
 DB: 4 Indels: 0
 Gaps: 0

US-09-770-509-2 (1-178) x US-09-134-001C-1996 (1-396)

QY 30 LeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeuGlyLysArg 49
 DB 392 ATGTGCGAATTAATTAACAGCTGATTAAGTGTGCTGCTGCTAATCTCAATTTGGAGACAAA 333
 QY 50 SerAlaGluGlnIleLysValAspIleGlnArgMetLeuGlnAspSerAsnMetLeuPhe 69
 DB 332 GCACAGAAAGATCAAGAGAACAAATTTGAAGACGCTATCCAGAGTCTGATATGTTTC 273
 QY 70 IleThrGlyMetGlyGlyGlyLysThrCysThrGlyAlaAlaProValAlaSerVal 89
 DB 272 GTAACGCTGGTATGCGGCGGCTGACAGAGGCTGACCGGCTTGTCTTAAATA 213
 QY 90 AlaArgLysLeuGlyIleLeuThrValGlyValAlaSerThrProPheArgSerGluGly 109
 DB 212 GCAGAAAGAAATGGGCTTAACTGTAGTGTGTTACCGCTCAATTCGTTTCAAGGT 153
 QY 110 ProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyrValAspThr 129
 DB 152 CGTAAGCGCCAAACACAGCGGACCTGCGCTAGATCTATGAAGCAGCAGTGTGATCA 93
 QY 130 LeuIleValAlaProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThrThrMetLeu 149
 DB 92 TTAATTTTATTCGCAAAAGATCGCTTATTAATATGATGCTGACAAATCTAGCCCAATGATG 33

RESULT 13

US-09-221-017B-658
 ; Sequence 658, Application US/09221017B
 ; Patent No. 6444799
 ; GENERAL INFORMATION:
 ; APPLICANT: Ross, Bruce C.
 ; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
 ; NUMBER OF SEQUENCES: 1120
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 PAGE MILL ROAD
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows
 ; SOFTWARE: FASTSEQ for Windows Version 2.0b
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/221,017B
 ; FILING DATE: 23-DEC-1998
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PP1182
 ; FILING DATE: 31-DEC-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PP1546
 ; FILING DATE: 30-JAN-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PP2911
 ; FILING DATE: 09-APR-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US98/01023
 ; FILING DATE: 10-DEC-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Monroy, Gladys H.
 ; REGISTRATION NUMBER: 32,430
 ; REFERENCE/DOCKET NUMBER: 27340-20021.00
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-813-5600

TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 658:
SEQUENCE CHARACTERISTICS:
LENGTH: 665 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...665
US-09-221-017B-658
Alignment Scores:
Pred. No.: 9.55e-38 Length: 665
Score: 369.50 Matches: 76
Percent Similarity: 67.92% Conservative: 32
Best Local Similarity: 47.80% Mismatches: 48
Query Match: 41.75% Indels: 3
DB: 4 Gaps: 2
US-09-770-509-2 (1-178) x US-09-221-017B-658 (1-665)
QY 3 GlnLeuGluGlyValAlaGluPheIleValAlaAsnThrAspGlyLeuValArgSer 22
DB 183 AAGGTGCGGAGTATCGTTCCTCTGCTGTATACCATGTCGACGCTCGACCGAAGC 242
QY 23 LeuAlaProValAlaValAspThrLeuGlyLysAspIleThrLysGlyAlaGlyLys 42
DB 243 GAGGTACCTGATCGGCTGGTACTCGGCGCTGAGGTGACCAAGCTCTGGGTCGGTAC 302
QY 43 LysProGluLeuGlyLysArgSerAlaGluGlnLysValAspIleGlnArgMetLeu 62
DB 303 CCGTCCGAAAGTGGCAGCAGCAGCTGCAGAGCCAGCCAGCCGACATACGCAATATG 362
QY 63 GlnAsp-----SerAsnMetLeuPheIleThrGlyGlyMetGlyGlyThrCysThr 80
DB 363 GATGATGCCCATACCGGTATGCTCTTCGTAAAGCGCCGCTATGGCGGAGCGGATACC 422
QY 81 GlyAlaAlaProValAlaValAspThrLeuGlyLysLeuValAlaGlyVal 100
DB 423 GGTGCGCCCTGTCTGATCGGTGAGTACGACGTGAATCATATCCTCACTGTGGCATC 482
QY 101 ValSerThrProPheArgSerGlyLysProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
DB 483 GTTACCATCCCATTCCTTTTCGAAAGCAGCCGCAAGATTCGACGCACTCGAAGGGGTG 542
QY 121 LysGluLeuAlaLysThrValAlaAspThrLeuIleValAlaProAsnGlnAsnLeuAla 140
DB 543 GAGGAAATCGCGCAAAACGATGATGCTTCGTCTGCTCAATATTAATGAACG--CTCCGC 599
QY 141 LeuAlaAspLysSerThrThrMetLeuGluAlaPheArgThrAlaAspValLeu 159
DB 600 ATCATATACAAGATCTTAAGCTCGACAAATGCTTTTCCCAAGCAGACGACACTG 656
RESULT 14
US-09-120-426-3
; Sequence 3, Application US/09120426
; Patent No. 6197300
; GENERAL INFORMATION:
; APPLICANT: Pueyo, Joanna Lynn
; APPLICANT: Lonetto, Michael A.
; TITLE OF INVENTION: ftsZ
; FILE REFERENCE: GMI0068
; CURRENT APPLICATION NUMBER: US/09/120,426
; FILING DATE: 1998-07-22
; EARLIER APPLICATION NUMBER: 60/055,720

EARLIER FILING DATE: 1997-08-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(669)
US-09-120-426-3
Alignment Scores:
Pred. No.: 2.82e-25 Length: 669
Score: 271.00 Matches: 52
Percent Similarity: 74.23% Conservative: 20
Best Local Similarity: 53.61% Mismatches: 25
Query Match: 30.62% Indels: 0
DB: 4 Gaps: 0
US-09-770-509-2 (1-178) x US-09-120-426-3 (1-669)
QY 82 AlaAlaProValAlaValAspThrLeuGlyLysLeuValAlaGlyVal 101
DB 1 TCTGCTCGTATGCTGTATCGCCCAAGATTTAGTGCCTGACAGTGTGTTGTA 60
QY 102 SerThrProPheArgSerGlyLysProAsnArgThrArgLeuAlaAsnAlaGlyVal 121
DB 61 ACAGCTCCCTTTGCTTTTAAAGAACTAAGCGTGCACAAATTTCTGTAAAGAAATCAT 120
QY 122 GlnLeuAlaLysThrValAlaAspThrLeuIleValAlaProAsnGlnAsnLeuAla 141
DB 121 CAACCTCGGACAGCATGATACACTATGATGATATCAACAACAAATTTGCTGAAT 180
QY 142 AlaAspLysSerThrThrMetLeuGluAlaPheArgThrAlaAspValLeuGlu 161
DB 181 GTTGATTAAGAAACACCGCTTTTGAGGCTCTTAGCGAAGCGGATTAACGTTCTCGTCA 240
QY 162 GlyAlaLysGlyValAlaPheAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
DB 241 GGTGTTCAAGGATTAACCATTTGATTACCAATCCAGATTAATTAACCTT 291
RESULT 15
US-08-961-083-123
; Sequence 123, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2

DB:	9	Gaps:	0
US-09-770-509-2 (1-178) x US-09-770-509-1 (1-535)			
OY	1	AlaSerGlnLeuGluGlyValGlnPheIleValAlaAsnThrAspCysGlnAlaLeuGly	20
Db	2	GCCTGCCAATTGGGAAGGTGTGGATTCAATTAGCCACACAGACTGTCCAGGCTTGGGA	61
OY	21	ArgSerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAla	40
Db	62	CGCTCGCTGGCCCCGACACAGATTCACGCTGGGGAAGAAATATCCCAAGGAGCTAGAGCT	121
OY	41	GlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArg	60
Db	122	GGATCCAAACCTGAGACTGGGTAAAGCCTCTGGCGAAGACAGACAGAAATGGATTCCAAAGC	181
OY	61	MetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyThrCysThr	80
Db	182	ATGTTACAGGAGACGACAACTGCTGTTTATCACGGGGGAATGGCGCGGCAACTCCACA	241
OY	81	GlyAlaAlaProValAlaAlaSerValAlaArgGluLeuGlyLeuThrValGlyVal	100
Db	242	GGAGCGGACCTGTGTGTGGCCAGTGTAGCCAGGAGACTGGGATCTTAACGTCGAGTA	301
OY	101	ValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal	120
Db	302	GTAACACACCGTCCGATCCGAAAGACCAATCGACATCGTCTGGCCAAATGCTGGAGTA	361
OY	121	LysGluLeuAlaLysTyrValAspThrLeuIleValAlaProAsnGlnAsnLeuAla	140
Db	362	AAAGACATGGCCAAATACCTGACACACTTAATTTGCTGCTCCACACAGAACTGTGGCT	421
OY	141	LeuAlaAspLysSerThrThrMetLeuGlnAlaPheArgTyrAlaAspAspValLeu	160
Db	422	TTGGCAGACAAAGACGACACCACTGTGGAAAGCTTCCGATATGCCACGACGCTGCTT	481
OY	161	GluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuAlaLeu	178
Db	482	GAGGAGATTAAAGTGTCAACGACTTGATCTTCCGCCGGACTTATCAATTGG	535
RESULT 2			
US-09-974-300-4664			
Sequence 4664, Application US/09974300			
Patent No. US20020146721A1			
GENERAL INFORMATION:			
APPLICANT: Berke, Randy M.			
APPLICANT: Clausen, ID Groth			
TITLE OF INVENTION: Methods For Monitoring Multiple Gene			
TITLE OF INVENTION: Expression			
FILE REFERENCE: 10085, 500-US			
CURRENT APPLICATION NUMBER: US/09/974, 300			
CURRENT FILING DATE: 2001-10-05			
PRIOR APPLICATION NUMBER: 09/680, 598			
PRIOR FILING DATE: 2000-10-06			
PRIOR APPLICATION NUMBER: 60/279, 526			
PRIOR FILING DATE: 2001-03-27			
NUMBER OF SEQ ID NOS: 8481			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 4664			
LENGTH: 950			
TYPE: DNA			
ORGANISM: Bacillus. clausii			
US-09-974-300-4664			
Alignment Scores:			
Pred. No.: 1.57e-57		Length: 950	
Score: 509.00		Matches: 101	
Percent Similarity: 74.86%		Conservative: 30	
Best Local Similarity: 57.71%		Mismatches: 44	
Query Match: 57.51%		Indels: 0	
DB: 10		Gaps: 0	
US-09-770-509-2 (1-178) x US-09-974-300-4664 (1-950)			

```

OY      4 LeuGluGlyValGluPheIleValAlaSerThrAspCysGlnAlaIleuGlyArgSerLeu 23
         |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      103 CTGCAAGGTGTTGATTATTCAGCATCATACAGATGCACACAGCATCCATTATCAAAA 162
OY      24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
         |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      163 GCGGAAAAGAGTTCCAACTTGCGCGCAAACTTACCGCGCGCTTGCGCTTGCGCGCAAC 222
OY      44 ProGluLeuGlyLysArgSerIleGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
         |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      223 CCAGAAATAGGCAAGAAAGCGGACAGGAGGAAAGCAACAGCTGGAAAGATTCTTACA 282
OY      64 AspSerAsnMetLeuPheIleThrGlyIleMetGlyGlyIleThrCysThrGlyAlaIle 83
         |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      283 GGTTCGTATGTCGCTTTATTTACTGCGCGAATGGGAGGAGAAACAGCACCGGAGCAGCT 342
OY      84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
         |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      343 CCGTCATCTTGCGAAGTGCAGAAAGACTTGGCGCATTCAGCGTCCGTTGTGACACGC 402
OY      104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
         |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      403 CCTTCTCTTTTGAAGGCGGTAGCGGCAAAACCAAGCATCTCTGGGATTCGACGCGTTA 462
OY      124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuAlaIleuAlaAsp 143
         |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      463 AAAGAAAAGGTGGACACATGTATCGTGATCCAAATGACCGGCTTATTGGAAATGTTGAT 522
OY      144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
         |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      523 AAAAATACGCGCATCTTGAAACGCTTCCGCGAGGAGGATTAAGTATTTGGCCAAAGAGTT 582
OY      164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
         |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      583 CAAGGATTCGCGATTATTTCCACAGCGCTGGGTGATCAACTT 627

RESULT 3
US-09-974-300-180
: Sequence 180, Application US/09974300
: Patent No. US20020146721A1
: GENERAL INFORMATION:
: APPLICANT: Beika, Randy M.
: APPLICANT: Clausen, ID Groth
: TITLE OF INVENTION: Methods For Monitoring Multiple Gene
: TITLE OF INVENTION: Expression
: FILE REFERENCE: 10085, 500-US
: CURRENT APPLICATION NUMBER: US/09/974, 300
: CURRENT FILING DATE: 2001-10-05
: PRIOR APPLICATION NUMBER: 09/680, 598
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 60/279, 526
: PRIOR FILING DATE: 2001-03-27
: NUMBER OF SEQ ID NOS: 8481
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 180
: LENGTH: 1134
: TYPE: DNA
: ORGANISM: Bacillus licheniformis
US-09-974-300-180

Alignment Scores:
Pred. NO.:      1.7e-56      Length:      1134
Score:          502.00      Matches:      97
Percent Similarity: 74.01%      Conservative: 34
Best Local Similarity: 54.80%      Mismatches: 46
Query Match:      56.72%      Indels:      0
DB:              10      Gaps:      0

US-09-770-509-2 (1-178) x US-09-974-300-180 (1-1134)
OY      2 SerGlnLeuGluGlyValGluPheIleValAlaSerThrAspCysGlnAlaIleuGlyArg 21
         |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```



```
Db 97 AATGACGCTTCAGGAGCGAGTCAGTTATCCAGTCAACAGCAGATGCTCAGGCTCTCAACCTG 156
Oy 22 SerLeuAlaProHISLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
Db 157 TCAAAACCGGAAACGAAATCAGATCGGTGCGAAGCTGCGGCGGCTGCGGCGCGGA 216
Oy 42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
Db 217 GCCAATCCGGAAAGTGGGCAAAAACCCGACAGGAAGAAAGAAACAATGAAAGAACG 276
Oy 62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyLysThrCysThrGly 81
Db 277 CTGAAAGCTGCAGATATGTTGTTCTCTCACAGCCGGAATGGCGGCGGCAACAGGAAGCGG 336
Oy 82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValAla 101
Db 337 GCGGACCTGTCATCGCAAAATCGCAAAAGATCTGGGCGCATGTGCTGCGGCTGTC 396
Oy 102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
Db 397 ACAAGCGCGTTTACCTTGAAGAGAAAGAAAGACAGCTTCAGGCTCAGCGGCTATTTC 456
Oy 122 GluLeuAlaLysTrpValAspThrLeuIleValValProAsnGlnAsnLeuAlaLeu 141
Db 457 GCATGGAAGGAAGCGGTGACACCTGATCGTCAATCCGAATGACCGCTCTTGAATATC 516
Oy 142 AlaAspLysSerThrThrMetLeuGlnAlaPheArgTrpValAlaAspValLeuLeuGlu 161
Db 517 GTCATTAATAACACACCGATCTGAAACCGTCCGGAAGGACACAGCTTCGCCGCA 576
Oy 162 GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
Db 577 GGTGTCAGGCGCATTTGACACTGATCCGAACGCGCTGACGTATCAACTT 627

RESULT 4
US-09-815-242-6542
; Sequence 6542, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6542
; LENGTH: 1233
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (1)...(1233)
; US-09-815-242-6542

Alignment Scores:
Pred. No.: 1,66-55 Length: 1233
Score: 495.00 Matches: 100
Percent Similarity: 73.14% Conservative: 28
Best Local Similarity: 57.14% Mismatches: 47
Query Match: 55.93% Indels: 0
Db: 10 Gaps: 0

US-09-770-509-2 (1-178) x US-09-815-242-6542 (1-1233)

Oy 4 LeuGluGlyValGluPheIleValAlaSerThrAspCysGlnAlaLeuGlyArgSerLeu 23
Db 106 GTTAAAGGGGTGAATTTATACAGCCAAATACAGACCTTCAGACATTAAACATTCAAAA 165
Oy 24 AlaProHISLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
Db 166 GCAGAAACAGTGTATTCATTAGCCCTCAATATACACTGCTGTAGCGCGCTTACAA 225
Oy 44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
Db 226 CCGAAGTGGCCCAAAACCTGCAAGAAAGTGAACAAGTTCAGATTCAGATTCATACAA 285
Oy 64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyLysThrCysThrGlyValAla 83
Db 286 GCGCGGATATATTTTCAATTTACTGCTGTATGGGTGGCGAAGTGTACAGTGTCTCG 345
Oy 84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValAlaSerThr 103
Db 346 CCAGTAGTTCGCAAAATCCCTAAGATTTAGCGCTTTAACAGTGTGCTAGTACTGTCT 405
Oy 104 PropheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
Db 406 CCATTAGTTCGTAAGATCCAAACCGTCTGTTCCCGCTGCAAGGATTCCTTATTA 465
Oy 124 AlaLysTrpValAspThrLeuIleValValProAsnGlnAsnLeuAlaLeuAlaAsp 143
Db 466 AAAGAAACGTTGATACACTATTAATATCTCAACAAACCGCTTATTAAGAAAGCTTAC 525
Oy 144 LysSerThrThrMetLeuGlnAlaPheArgTrpValAlaAspValLeuLeuGluGlyVal 163
Db 526 AAGAAACCGCAATGCTTGAAGCATTTAGAGAGCTGATATATGATTAACGTCAGGTCT 585
Oy 164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
Db 586 CAAGGATTCAGATTTAATCACTGCACCGAGGTACGTAACCTTA 630

RESULT 5
US-09-815-242-3854
; Sequence 3854, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
```

PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242, 578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253, 625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257, 931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269, 308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3854
LENGTH: 1239
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-815-242-3854

Alignment Scores:
Pred. No.: 1,61e-55 Length: 1239
Score: 495.00 Matches: 100
Percent Similarity: 73.14% Conservative: 28
Best Local Similarity: 57.14% Mismatches: 47
Query Match: 55.93% Indels: 0
DB: 10 Gaps: 0

US-09-770-509-2 (1-178) x US-09-815-242-3854 (1-1239)

QY 4 LeuGlulGlyValAluPheIleValAlaAnThrAspCysGlnAlaLeuGlyArgSerLeu 23
DB 106 GTTAAAGCGGTGAATTTATACAGCAATACAGCGTTCAAGATTAAACATTCAAAA 165
QY 24 AlaProHisIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
DB 166 GCAGAAACAGGTGATTAATTAAGGCGCTTAATACACTGCTGTTAGTCCCGTACACA 225
QY 44 ProGluLeuGlyLysArgSerAlaGluGlnLysValAspIleGlnArgMetLeuGln 63
DB 226 CCGTAAGTTGGCCAAAAGCTCCAGAAAGAAAGTACAGATTCAGATTCATTACAA 285
QY 64 AspSerAsnMetLeuPheIleThrGlyMetGlyGlyThrCysThrGlyAlaAla 83
DB 286 GCGCGCATATGATTTTCTACTGCTGATGGTGGCGAAGTCAAGTCAAGTCTCGC 345
QY 84 ProValAlaIleSerValAlaArgGluLeuGlyIleThrValGlyValSerThr 103
DB 346 CCAGTAGTTCAAAAGTAAAGATTAAGTGGCTTAAAGTGTGTAGTACTCGT 405
QY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
DB 406 CCATTAGTTTGAAGTCCAAAAGCTGCTGTTTCCCGCTGAAGGATTCCTTATTA 465
QY 124 AlaLysTyrValAspThrLeuIleValAlaProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
DB 466 AAAGAAACGTTGATACATTAATTAATCTCAAAACACCGCTTATTAGAAGTCTGTAC 525
QY 144 LysSerThrMetLeuGlnAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
DB 536 AAGAAACCCCAATGCTTAAGCATTTAGAGAGCTGATATATATTACGTCAGGTGT 585
QY 164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
DB 586 CAAGGATTTGATTTAATCACTGCACCCAGTTAGCTTAACCTTA 630

RESULT 6
US-09-070-927A-128/c
Sequence 128, Application US/09070927A
Patent No. US20020120116A1
GENERAL INFORMATION:
APPLICANT: Charles A. Kunsch
Patrick J. Dillon
Steven Barash
TITLE OF INVENTION: Enterococcus faecalis polynucleotides and polypeptides
NUMBER OF SEQUENCES: 982

CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070, 927A
FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046, 655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044, 031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066, 009
FILING DATE: 1997-11-14

ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 128:
SEQUENCE CHARACTERISTICS:
LENGTH: 32768 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 128:
US-09-070-927A-128

Alignment Scores:
Pred. No.: 1,81e-53 Length: 32768
Score: 495.00 Matches: 100
Percent Similarity: 73.14% Conservative: 28
Best Local Similarity: 57.14% Mismatches: 47
Query Match: 55.93% Indels: 0
DB: 10 Gaps: 0

US-09-770-509-2 (1-178) x US-09-070-927A-128 (1-32768)

QY 4 LeuGlulGlyValAluPheIleValAlaAnThrAspCysGlnAlaLeuGlyArgSerLeu 23
DB 27971 GTTAAAGCGGTGAATTTATACAGCAATACAGCGTTCAAGATTAAACATTCAAAA 27912
QY 24 AlaProHisIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
DB 27911 GCAGAAACAGGTGATTAATTAAGGCGCTTAATACACTGCTGTTAGTCCCGTACACA 27852
QY 44 ProGluLeuGlyLysArgSerAlaGluGlnLysValAspIleGlnArgMetLeuGln 63
DB 27851 CCGTAAGTTGGCCAAAAGCTCCAGAAAGAAAGTACAGATTCAGATTCATTACAA 27792
QY 64 AspSerAsnMetLeuPheIleThrGlyMetGlyGlyThrCysThrGlyAlaAla 83
DB 27791 GCGCGCATATGATTTTCTACTGCTGATGGTGGCGAAGTCAAGTCTCGC 27732
QY 84 ProValAlaIleSerValAlaArgGluLeuGlyIleThrValGlyValSerThr 103
DB 27731 CCAGTAGTTCAAAAGTAAAGATTAAGTGGCTTAAAGTGTGTAGTACTCGT 27672
QY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
DB 27671 CCATTAGTTTGAAGTCCAAAAGCTGCTGTTGCTGTGAAGGATTCCTTATTA 27612


```

Db      626 GGTGACGACCCCTCATGCTGCAGGCTGCCTTAGATGCTGTATCCTACCGTAGCTGTC 685
Qy      101 ValSerThrProheargSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
Db      686 GTTACTACAGCCGCTCCGCTTGAAGGAAACACCGCTGCAGACCTTCCGACACAGGCTTC 745
Qy      121 LysGluLeuAlaLysTyrValAspThrLeuIleValAlaProAsnGlnAsnLeuAla 140
Db      746 GCTGAACGTAGAGGATAGCGTGCATACGATGCTGTGATCCGGAACCAAACTTGTTACAC 805
Qy      141 LeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeu 160
Db      806 ATGTCAAAATGAGGCGACCTGCTGATGAGCAGCATTCAGAAATGGCGACAAATGCTTCG 865
Qy      161 GluGlyValLysGlyValThrAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
Db      866 GACGGGTCAAGACACTTCGATTTGATGTGATGCTGCGCTCATTACCTT 919

```

RESULT 9

```

US-09-815-242-4715
; Sequence 4715, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4715
; LENGTH: 1172
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4715

```

Alignment Scores:

Pred. No.:	1,41e-53	Length:	1172
Score:	480.00	Matches:	95
Percent Similarity:	73.148	Conservative:	33
Best Local Similarity:	54.298	Mismatches:	47
Query Match:	54.248	Indels:	0
DB:	10	Gaps:	0

US-09-770-509-2 (1-178) x US-09-815-242-4715 (1-1172)

```

Qy      4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
Db      103 ATGAATATATGTTGAATTTATCGCTATCAACACAGACGGCTTAAGCTTAACCTATCTAAA 162

```

```

Qy      24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
Db      163 GCTGAATCTAAATCAATCAATGCTGTAATAATTAACACGCTGCTTAGACGAGCAAGTAAT 222
Qy      44 ProLiuLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
Db      223 CCTAATAATCGGTAAATAAAGCTGCAAGAGAAATCTCGGAACAAATTAAGATGCAATCCAA 282
Qy      64 AsperAsnMetLeuPheIleThrGlyGlyMetGlyGlyThrCysThrGlyAlaAla 83
Db      283 GGTGACAGCATGATATTGTTGCTTCTGATGCGTGGCGGCAACCTGCTGCTGACGA 342
Qy      84 ProValAlaLysSerValAlaArgGluLeuGlyIleLeuThrValGlyValSerThr 103
Db      343 CCACTCGTTGCTAAATATTCAGAAAGAAATGGCGCATTAACGTGTGTTGTAACCTG 402
Qy      104 ProheargSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
Db      403 CCATTTAGTTTGAAGGACGTAAACGTCAAACTCAAGCTGCTGAGTAGAAGCTATG 462
Qy      124 AlaLysTyrValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaAsp 143
Db      463 AAAGCTGACATGATACATTAATGCTTATACCAATGACCGTTTATGATATCGTTGAC 522
Qy      144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
Db      523 AAATCTAGCGCAATGATGAGACATTTAAAGAACTGACCAACGCTGTACGCCAAGTGTA 582
Qy      164 LysGlyValThrAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
Db      583 CAAGGTATCTCAGACTTAATCGCTGTTCTGATGAAGTAACCTA 627

```

RESULT 10

```

US-09-815-242-8625
; Sequence 8625, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8625
; LENGTH: 1179
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1179)

```

US-09-815-242-8625

Alignment Scores:

Pred. No.:	1,42e-53	Length:	1179
Score:	480.00	Matches:	95
Percent Similarity:	73.14%	Conservative:	33
Best Local Similarity:	54.29%	Mismatches:	47
Query Match:	54.24%	Indels:	0
DB:	10	Gaps:	0

US-09-770-509-2 (1-178) x US-09-815-242-8625 (1-1179)

```

OY 4 LeuGlUGlyValAlaGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 103 ATGAATTAATGTTGAATTTATCGTATCAACACAGCGCTCAAGCTTAACTTAACTTAA 162
OY 24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 163 GCTGAATCTAAATCAAAATCGTGGAATAAATTACACCTGGTTAAGGACAGACCTTAT 222
OY 44 ProGluLeuGlyLysArgSerAlaGluGlnLysValAspIleGlnArgMetLeuGln 63
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 223 CCGTAATCGGTAAATAAAGCGCAGAGCAATCTCGTGAACAATTGAAGATGCAATCCAA 282
OY 64 AspSerAsnMetLeuPheIleThrGlyLysMetGlyGlyThrCysThrGlyAlaAla 83
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 283 GGTGACACATGATTTGTTACTTCTGATGGGTGCGGAACGTGATCTGTCACAGA 342
OY 84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValAlaSerThr 103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 343 CCGATCTCTTAAATTCGAAAGAAATGGGCGCATTAACGTGGTGTCTGTAACCTGCT 402
OY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 403 CCAATTGTTTGAAGAGACGTAACGTAACGTAACGCTGCTGCTGAGAGTACAGACTATG 462
OY 124 AlaLysTyValAlaAspThrLeuIleValAlaProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 463 AAGCTGCATGATGATTAATGCTTATACCAATGACCGTTTATGATTCGTTGAC 522
OY 144 LysSerThrThrMetLeuGluAlaPheArgTyrrAlaAspAspValLeuLeuGluGlyVal 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 523 AAATCTACGCCAATGATGAGGACATTTAAAGAACGTCGCAACGTCGTCGCCAAGGTGA 582
OY 164 LysGlyValThrAspLeuIleValAlaArgProGlyLeuLeuAsnLeu 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 583 CAAGGTATCTCAGACTTAATCGCTGTTCTGCTGTAAGTAAACTTA 627

```

RESULT 11

```

US-08-781-986A-342
; Sequence 342, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248PP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 342:

SEQUENCE CHARACTERISTICS:

LENGTH: 3931 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-781-986A-342

Alignment Scores:

Pred. No.:	8.05e-53	Length:	3931
Score:	480.00	Matches:	95
Percent Similarity:	73.14%	Conservative:	33
Best Local Similarity:	54.29%	Mismatches:	47
Query Match:	54.24%	Indels:	0
DB:	7	Gaps:	0

US-09-770-509-2 (1-178) x US-08-781-986A-342 (1-3931)

```

OY 4 LeuGlUGlyValAlaGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 684 ATGAATTAATGTTGAATTTATCGTATCAACACAGCGCTCAAGCTTAACTTAACTTAA 743
OY 24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 744 CCGATCTCTTAAATTCGAAAGAAATGGGCGCATTAACGTGGTGTCTGTAACCTGCT 803
OY 44 ProGluLeuGlyLysArgSerAlaGluGlnLysValAspIleGlnArgMetLeuGln 63
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 804 CCGTAATCGGTAAATAAAGCGCAGAGCAATCTCGTGAACAATTGAAGATGCAATCCAA 863
OY 64 AspSerAsnMetLeuPheIleThrGlyLysMetGlyGlyThrCysThrGlyAlaAla 83
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 864 GGTGACACATGATTTGTTACTTCTGATGGGTGCGGAACGTGATCTGTCACAGA 923
OY 84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValAlaSerThr 103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 924 CCGATCTCTTAAATTCGAAAGAAATGGGCGCATTAACGTGGTGTCTGTAACCTGCT 983
OY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 984 CCAATTGTTTGAAGAGACGTAACGTAACGTAACGCTGCTGCTGAGAGTACAGACTATG 1043
OY 124 AlaLysTyValAlaAspThrLeuIleValAlaProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1044 AAGCTGCATGATGATTAATGCTTATACCAATGACCGTTTATGATTCGTTGAC 1103
OY 144 LysSerThrThrMetLeuGluAlaPheArgTyrrAlaAspAspValLeuLeuGluGlyVal 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1104 AAATCTACGCCAATGATGAGGACATTTAAAGAACGTCGCAACGTCGTCGCCAAGGTGA 1163
OY 164 LysGlyValThrAspLeuIleValAlaArgProGlyLeuLeuAsnLeu 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1164 CAAGGTATCTCAGACTTAATCGCTGTTCTGCTGTAAGTAAACTTA 1208

```

RESULT 12

```

US-09-754-608-1
; Sequence 1, Application US/09754608
; Patent No. US20020004580A1
; GENERAL INFORMATION:
; APPLICANT: Fueno, Joanna Lynn
; APPLICANT: Lonetto, Michael A.
; APPLICANT: Pearce, Kenneth
; TITLE OF INVENTION: fts2
; FILE REFERENCE: GM10068

```


OY 166 ValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
DB 592 ATACCGATTGATTACCAATCCAGATTGATTACCTT 630
RESULT 14
US-09-738-626-2360
Sequence 2360, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, MAKOTO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738, 626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 2360
LENGTH: 1326
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
US-09-738-626-2360
Alignment Scores:
Pred. No.: 3,49e-52 Length: 1326
Score: 470.00 Matches: 92
Percent Similarity: 73.14% Conservative: 36
Best Local Similarity: 52.57% Mismatches: 47
Query Match: 53.11% Indels: 0
Gaps: 0
DB: 9
US-09-770-509-2 (1-178) x US-09-738-626-2360 (1-1326)
OY 4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
DB 94 CTCGAAGCGGTGAGTTCATCGCGGTGAACACGACGCTCATGTCCTCAT 153
OY 24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
DB 154 GCCGACGTAAAGCTCATATCGGACGTGAAGCTACCGGTCTTGGTGGCGGCAGAC 213
OY 44 ProGluLeuGlyLysArgSerAlaGluGlnLysValAspIleGlnArgMetLeuGln 63
DB 214 CCAGAGTTGGAGCTGCTCGCGCAGAGATCACAAGAACGAATCGAAGAACCATCAG 273
OY 64 AspSerAsnMetLeuPheIleThrGlyLysMetGlyLysGlyThrCysThrGlyAlaAla 83
DB 274 GCGCGCATATGCTCTTCTTACCGCGCGCAAGGTGTGGCAGCACTGTGTCTCA 333
OY 84 ProValAlaIleSerValAlaArgGluLeuGlyIleLeuThrValGlyValAlaSerThr 103
DB 334 CCAAGTGTGGCAGGATCGCCAGAAAGATGGCGCACTGACCATGTGTGTGACCAAG 393
OY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValGlySerLeu 123
DB 394 CATTTCAGATGCAAGGCGCTCGCCGACCTCCGACAGCAGAAAGAGCATCGACACTG 453
OY 124 AlaLysTyrValAlaSerThrLeuIleValAlaProAsnGlnAsnLeuLeuAlaAlaSer 143

DB 454 AAGAGCTGTGCAGACCCCTCATCTATTTCAAAGACCGCTGTTGAGCTGGCGCAT 513
OY 144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValIleLeuGluGlyAla 163
DB 514 GCGAACCTGTCTCATATGGAAGCGTTCGCGCAGCGCATGATTCACCAATGTGT 573
OY 164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
DB 574 CAGGTTATTACCAACTGATCACCATTCCCTGTGTGATCAACGTG 618
RESULT 15
US-09-738-626-1/c
Sequence 1, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, MAKOTO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738, 626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 1
LENGTH: 3309400
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
US-09-738-626-1
Alignment Scores:
Pred. No.: 2,74e-47 Length: 3309400
Score: 470.00 Matches: 92
Percent Similarity: 73.14% Conservative: 36
Best Local Similarity: 52.57% Mismatches: 47
Query Match: 53.11% Indels: 0
Gaps: 0
DB: 9
US-09-770-509-2 (1-178) x US-09-738-626-1 (1-3309400)
OY 4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
DB 2280122 CTCGAAGCGGTGAGTTCATCGCGGTGAACACGACGCTCATGTCCTCAT 2280063
OY 24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
DB 2280062 GCGGACGTAAAGCTCATATCGGACGTGAAGCTACCGGTCTTGGTGGCGGCAGAC 2280003
OY 44 ProGluLeuGlyLysArgSerAlaGluGlnLysValAspIleGlnArgMetLeuGln 63
DB 2280002 CCAGAGTTGGAGCTGCTCGCGCAGAGATCACAAGAACGAATCGAAGAACCATCAG 2279943
OY 64 AspSerAsnMetLeuPheIleThrGlyLysMetGlyLysGlyThrCysThrGlyAlaAla 83
DB 2279942 GCGCGCATATGCTCTTCTTACCGCGCGCAAGGTGTGGCAGCACTGTGTCTCA 2279883
OY 84 ProValAlaIleSerValAlaArgGluLeuGlyIleLeuThrValGlyValAlaSerThr 103
DB 2279882 CCAAGTGTGGCAGGATCGCCAGAAAGATGGCGCACTGACCATGTGTGTGACCAAG 2279823

```
OY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValIysGluLeu 123
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2279822 CCTTCGAGTTCGAGAAGCCGTCGCCGCTACTCGCCAGGCGAGAAGAAGCATCGCAGCAGCTG 2279763
OY 124 AlaLysThrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2279762 AAGGAGGCTGCGCACACCCCTCATCGTATTCCAAACGACCGCCCTGCTGAGCTGGGCGAT 2279703
OY 144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspValIleLeuGluGlyVal 163
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2279702 GCGAACCTGTCATCATGAGCGCTTCGCGCAGCCGATGAGTCTCCACATGATGTT 2279643
OY 164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 2279642 CAGGGTATTACCAACCTGATCACCATCCCTGTGTGATCAACGTG 2279598
```

Search completed: June 2, 2003, 13:17:12
Job time : 7750 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 2, 2003, 08:52:57 ; Search time 2738 Seconds
(without alignments)
1634.537 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGEVIVANTDCALG.....LLEGKVTDLIVRGLINL 178

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-q/cgn2_1/USOTO.spool/US09770509/runat_27052003_083651_18863/app.query.fasta_1.327
-DB=Pending_Patents_NA_Main -QMT=fastap -SUFFIX=trnmp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=p2n -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09770509_083651_1.2950_0runat_27052003_083651_18863 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Pending_Patents_NA_Main:*
2: /cgn2_6/ptodata/1/pna/PCRUS_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US08_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US095_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US095E_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095F_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*

29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US6023_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	885	100.0	535	30	US-09-770-509-1
2	550	62.1	1473	38	US-10-015-127-5567
3	550	62.1	24740	38	US-10-015-127-5522
4	509	57.5	950	36	US-09-974-300-4664
5	509	57.5	2074	71	US-60-279-526-159
6	506	57.2	111309	1	PCT-US98-12764-3

7	506	57.2	111309	32	US-09-830-228-3	Sequence 3, Appl 1
8	506	57.2	910715	1	PCT-US98-12764-1	Sequence 1, Appl 1
9	506	57.2	910715	32	US-09-830-228-1	Sequence 1, Appl 1
10	505	57.1	7375	26	US-09-663-779-11189	Sequence 1189, Ap
11	503	56.8	1864	29	US-09-754-468-42	Sequence 42, Appl
12	502	56.7	1134	36	US-09-974-500-180	Sequence 180, App
13	496	56.0	3993	48	US-60-045-649-854	Sequence 854, App
14	496	56.0	3993	48	US-60-046-653-853	Sequence 853, App
15	495	55.9	1233	1	PCT-US02-03887-6542	Sequence 6542, Ap
16	495	55.9	1233	31	US-09-815-242-6542	Sequence 6542, Ap
17	495	55.9	1233	39	US-10-072-851-6542	Sequence 6542, Ap
18	495	55.9	1239	1	PCT-US02-03887-3854	Sequence 3854, Ap
19	495	55.9	1239	31	US-09-815-242-3854	Sequence 3854, Ap
20	495	55.9	1239	39	US-10-072-851-3854	Sequence 3854, Ap
21	495	55.9	1308	15	US-09-134-000-1253	Sequence 1253, Ap
22	495	55.9	3904	50	US-60-068-217-757	Sequence 757, App
23	495	55.9	32768	14	US-09-070-927-127	Sequence 128, App
24	495	55.9	32768	14	US-09-070-927-128	Sequence 128, App
25	494	55.8	1245	15	US-09-107-532-3641	Sequence 3641, App
26	494	55.8	1245	15	US-09-107-532a-3641	Sequence 3641, App
27	494	55.8	13968	49	US-60-050-444-716	Sequence 716, App
28	494	55.8	13968	50	US-60-068-186-720	Sequence 720, App
29	485	54.8	1140	28	US-09-712-363-75	Sequence 75, Appl
30	485	54.8	37586	29	US-09-754-168-27	Sequence 27, Appl
31	482	54.5	1662	19	US-09-514-000-5249	Sequence 5249, App
32	482	54.5	1752	29	US-09-739-449-3270	Sequence 3270, App
33	482	54.5	1752	31	US-09-803-110-3270	Sequence 3270, App
34	482	54.5	1857	60	US-60-168-139-330	Sequence 1162, App
35	482	54.5	27239	19	US-09-514-000-330	Sequence 330, App
36	482	54.5	184668	29	US-09-739-449-207	Sequence 207, App
37	482	54.5	184668	31	US-09-803-110-207	Sequence 207, App
38	481	54.4	1433	30	US-09-770-609-9	Sequence 9, Appl 1
39	480	54.2	1158	14	US-09-037-934-12	Sequence 12, Appl 1
40	480	54.2	1159	24	US-09-620-008-1462	Sequence 1462, App
41	480	54.2	1159	58	US-60-144-883-1462	Sequence 1462, App
42	480	54.2	1170	16	US-09-266-541-1	Sequence 1, Appl 1
43	480	54.2	1172	1	PCT-US02-03887-4715	Sequence 4715, Ap
44	480	54.2	1172	31	US-09-815-242-4715	Sequence 4715, Ap
45	480	54.2	1172	39	US-10-072-851-4715	Sequence 4715, Ap

ALIGNMENTS

```

RESULT 1
US-09-770-509-1
: Sequence 1, Application US/09770509
: GENERAL INFORMATION:
:   APPLICANT: Katagiri, F.
:   TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
:   TITLE OF INVENTION: ANTIMICROBIAL-SPECIFIC BIODES
:   FILE REFERENCE: NA011.018A
:   CURRENT APPLICATION NUMBER: US/09/770,509
:   CURRENT FILING DATE: 2001-01-25
:   NUMBER OF SEQ ID NOS: 32
:   SOFTWARE: FastSeq for Windows Version 4.0
:   SEQ ID NO 1
:   LENGTH: 535
:   TYPE: DNA
:   ORGANISM: Phytophthora infestans
:   FEATURE:
:   NAME/KEY: CDS
:   LOCATION: (2)...(535)
:   US-09-770-509-1

```

Alignment Scores:	
Pred. No.:	3,97e-84
Score:	885.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	30
Length:	535
Matches:	178
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-770-509-2 (1-178) x US-09-770-509-1 (1-535)

[illegible]

```

RESULT 2
; US-10-015-127-5607
; Sequence 5607, Application US/10015127
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley G.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Sphingomonas elodea genome sequences and uses thereof
; FILE REFERENCE: 38-10/1518062B
; CURRENT APPLICATION NUMBER: US/10/015.127
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/252,455
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 14357
; SEQ ID NO 5607
; LENGTH: 1473
; TYPE: DNA
; ORGANISM: Sphingomonas elodea
US-10-015-127-5607

```

Alignment Scores:	
Pred. No.:	5,61e-48
Score:	550.00
Percent Similarity:	79.10%
Best Local Similarity:	59.32%
Query Match:	6.15%
DB:	38
Length:	1
Matches:	1
Conservative:	3
Mismatches:	3
Indels:	0
Gaps:	0

US-09-770-509-2 (1-178) x US-10-015-127-5607 (1-1473)

[illegible]

```
Db 166 TCGGTGCGCCGACGCGATTCAGCTGGCGCGAAGATCAGCAGGCGCGCGCGCG 225
Qy SerIysProGluLeuGlyValArgSerAlaGluGlnGlnValAspIleGlnArgMet 61
Db 226 TCGCGCGCCCAATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 285
Qy LeuGlnAspSerAsnMetLeuPheIleThrGlyMetGlyGlyGlyThrCysThrGly 81
Db 286 CTCGAAGGCGCGCATGCTGCTTCATCGCGCGCGCGCGCGCGCGCGCGCGCG 345
Qy AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyVal 101
Db 346 GCGCGACCGGTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 405
Qy SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 121
Db 406 ACCAAGCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 465
Qy GlnLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuAlaLeu 141
Db 466 GAATCGCAGAAATGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 525
Qy AlaAspLysSerThrThrMetLeuGlnAlaPheArgTyrAlaAspAspValLeuGlu 161
Db 526 GCCAATGCGAAGCAGCGCTTCAGGAAAGCGTTCGAGATGCGCGCGCGCG 585
Qy GlnValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
Db 586 GGTGTGCGCGCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 636

RESULT 3
US-10-015-127-522/C
: Sequence 522, Application US/10015127
: GENERAL INFORMATION:
: APPLICANT: Bower, Stanley G.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: TITLE OF INVENTION: Sphingomonas elodea genome sequences and uses thereof
: FILE REFERENCE: 38-10(15806)B
: CURRENT APPLICATION NUMBER: US/10/015,127
: PRIORITY FILING DATE: 2001-10-29
: PRIOR APPLICATION NUMBER: US 60/252,455
: PRIORITY FILING DATE: 2000-11-22
: NUMBER OF SEQ ID NOS: 14357
: SEQ ID NO 522
: LENGTH: 24740
: TYPE: DNA
: ORGANISM: Sphingomonas elodea
US-10-015-127-522

Alignment Scores:
Pred. No.: 1,68e-46 Length: 24740
Score: 550.00 Matches: 105
Percent Similarity: 79.10% Conservative: 35
Best Local Similarity: 59.32% Mismatches: 37
Query Match: 62.15% Indels: 0
Gaps: 0

US-09-770-509-2 (1-178) x US-10-015-127-522 (1-24740)
Qy 2 SerGlnLeuGlnGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
Db 3486 GCGGAGGTGACAGGGGTGCGATTCGCTGCGCAACCGACGCGCGCGCGCG 3427
Qy 22 SerLeuAlaProHisIleThrLeuGlnLysAspIleThrLysGlyLeuGlyAlaGly 41
Db 3426 TCGGTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3367
Qy 42 SerLysProGluLeuGlnGlyValArgSerAlaGluGlnGlnValAspIleGlnArgMet 61
Db 3366 TCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3307
Qy 62 LeuGlnAspSerAsnMetLeuPheIleThrGlyMetGlyGlyGlyThrCysThrGly 81
```

```
Db 3306 CTCGAAGCGCGCGCATATGCTTCATCGCGCGCGCGCGCGCGCGCGCGCG 3247
Qy 82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyVal 101
Db 3246 GCGCGACCGGTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3187
Qy 102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 121
Db 3186 ACCAAGCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3127
Qy 122 GlnLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuAlaLeu 141
Db 3126 GAATCGCAGAAATGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3067
Qy 142 AlaAspLysSerThrThrMetLeuGlnAlaPheArgTyrAlaAspAspValLeuGlu 161
Db 3066 GCCAATGCGAAGCAGCGCTTCAGGAAAGCGTTCGAGATGCGCGCGCG 3007
Qy 162 GlnValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
Db 3006 GGTGTGCGCGCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2956

RESULT 4
US-09-974-300-4664
: Sequence 4664, Application US/09974300
: GENERAL INFORMATION:
: APPLICANT: Berka, Randy M.
: APPLICANT: Clausen, Id Groth
: TITLE OF INVENTION: Methods For Monitoring Multiple Gene
: FILE REFERENCE: 10085, 500-US
: CURRENT APPLICATION NUMBER: US/09/974,300
: PRIORITY FILING DATE: 2001-10-05
: PRIOR APPLICATION NUMBER: 09/680,598
: PRIORITY FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 60/279,526
: NUMBER OF SEQ ID NOS: 8481
: SOFTWARE: PastSeq for Windows Version 4.0
: SEQ ID NO 4664
: LENGTH: 950
: TYPE: DNA
: ORGANISM: Bacillus clausii
US-09-974-300-4664

Alignment Scores:
Pred. No.: 7.56e-44 Length: 950
Score: 509.00 Matches: 101
Percent Similarity: 74.86% Conservative: 30
Best Local Similarity: 57.71% Mismatches: 44
Query Match: 57.51% Indels: 0
Gaps: 0

US-09-770-509-2 (1-178) x US-09-974-300-4664 (1-950)
Qy 4 LeuGlnGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
Db 103 CTCGAAGGTGATGATTTATTCAGTCAATACAGATGACACACATTCATTATCAAAA 162
Qy 24 AlaProHisIleThrLeuGlnLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
Db 163 GCGGAAAGAAAGTTCACACTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 222
Qy 44 ProGluLeuGlnGlyValArgSerAlaGluGlnGlnValAspIleGlnArgMetLeuGln 63
Db 223 CCAGAAATAGCGAAGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 282
Qy 64 AspSerAsnMetLeuPheIleThrGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
Db 283 GGTTCGTGATGATGCTTTATTAATGCGCGAATGGAGAGAAACGCGCGCGCG 342
Qy 84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValSerThr 103
```

Db	343	CCTGTCATTCGTGGAAGTGGCAAAAGAGCTTGGCCGATTAACGGTCGGTGTGTGCACGC	402
QY	104	ProPheargSerGluGIYProasnArgThrArgLeuAlaAsnIacIyValLysGluLeu	123
Db	403	CGTTTCTCTTTGGAAGGCGGTAAAGGGGCAAAACCAAGCATCTCTGGGARTGCAGCGCTTA	462
QY	124	AlaIysTyrValAspThrLeuIleValValProAsnGlnAsnLeuAlaLeuAlaAsp	143
Db	463	AAAGCAAAAAGGGACACATGATCTGTGATCCAAATAGACCGCTATTGTGAAAGGGTTGAT	522
QY	144	LysSerThrThrMetLeuGluGlnAlaPheArgTyrAlaAspAspValLeuLeuGluGluVal	163
Db	523	AAAAATACGCCGATGCTTGAAGCGCTTCGCGCGAGGGGATAAAGTATATGGCCCAAGAGTT	582
QY	164	LysGluValThrAspLeuIleValArgProGluLeuIleAsnLeu	178
Db	583	CAAGGCATTCGCGATTATATGCGCACGCGCTGGGTGATCAACCTT	627

```

RESULT 5
US-60-279-526-159/c
: Sequence 159, Application US/60279526
: GENERAL INFORMATION:
: APPLICANT: Beika, Randy M.
: APPLICANT: Clausen, Ib Groth
: APPLICANT: Sorokline, Alexei
: APPLICANT: Apolotline, Alexandre
: APPLICANT: Lapidus, Alla
: TITLE OF INVENTION: Methods for Monitoring Multiple Gene
: FILE REFERENCE: 5963,000-US
: CURRENT APPLICATION NUMBER: US/60/279,526
: CURRENT FILING DATE: 2001-03-27
: NUMBER OF SEQ ID NOS: 2232
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 159
: LENGTH: 2074
: TYPE: DNA
: ORGANISM: B. clausii
US-60-279-526-159

```

Alignment Scores:		
Pred. No.:	1.94e-43	207
Score:	509.00	Matches: 101
Percent Similarity:	74.86%	Conservative: 30
Best Local Similarity:	57.71%	Mismatches: 44
Query Match:	57.51%	Indels: 0
DB:	71	Gaps: 0

QY	4	LeuSibuliyValGluPheIleValAlaSnTrpSCysGlnAlaLeuGlyArgSerLeu	23
Db	857	CTGCAAGGTTGGATTATTATGCACTACATACAGATGCACAAAGCACTCCATTATCAAAA	796
QY	24	AlaProHisIysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys	43
Db	797	GCGGAAAAAGAGTTGCCAATCTGGCGGCAAACTACTCTCGGCGCTTGCGCGTGCAC	738
QY	44	ProGluLeuGlyLysArgSerAlaGluGlnGlyValAspIleGlnArgMetLeuGln	63
Db	737	CCAGAAATTAGCGCAAGAAAAGCGCCAGAGAAACAGAAACAGCTGGAAAGATCTTACA	678
QY	64	AspSerAsnMetLeuPheIleThrGlyIleMetGlyGlyIleThrCysThrGlyAlaAla	83
Db	677	GGTTTCGATATGAGCTTTATTACTGCGCGAATGGGAGGAGAAACAGCACCGGAGCAGCT	618
QY	84	ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr	1030
Db	617	CCGTGCTATTCTCGAAGTCGCAAAAGAGCTTGCGCATGACGCTGGTGTTCGACACGC	558
QY	104	ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu	123

Db 557 CCTTCTCTTTTGAAGGGCCGTAAAGCGGCAAAACCAAGCATCTCTGGATTGACGGTTA 498

QY 124 ALALSYTYRVALASpThrIleuIleValIProAsrGlnAsnIleuIleuAlaIasp 143

Db 497 AAAGAAAAAGTGGACACATGTGTCGGATTCCAAATGACCGCTTTATTGGAAGTGTGAT 438

QY 144 LysSerThrThrMetIleuGlnIalaphaArgTYRAlaAspAspValIleuIleuGluIVal 163

Db 437 AAAAATACGCCGATGCTTGAAACGCTTCCCGCGAGCGGATTAATGTATTGGCCCAAGAGTT 378

QY 164 LysGluValThrAspIleuIleValAlaArgProGluIleuIleAsnIleu 178

Db 377 CAAGGCATTCGATTAATTATGCCACGCGCTGGGTGATCAACCTT 333

RESULT 6
 PCT-US98-12764-3
 Sequence 3, Application PC/TUS9812764
 GENERAL INFORMATION:
 APPLICANT: Gil Choi et. al.
 TITLE OF INVENTION: Borrelia burgdorferi Polynucleotides and Sequences
 NUMBER OF SEQUENCES: 155
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US98/12764
 FILING DATE: Herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Brookes, A. Anders
 REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: PB370PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 111309 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear

Primer Name:	4.92e-41	Length:	111309
Primer No.:	506.00	Matches:	101
Score:	74.42%	Conservative:	27
Percent Similarity:	58.72%	Mismatches:	44
Best Local Similarity:	57.18%	Indels:	0
Query Match:	1	Gaps:	0
DB:			
US-09-770-509-2 (1-178) x PCR-US98-12764-3 (1-111309)			
Qy	7	ValGluPheIleValAlaAspThrAspCysGlnAlaLeuGlyArgSerLeuAlaProHis	26
Db	12947	GTGTAATTATTGTGGCTAATACCGATCTTCAGGCTCTCCAAACTTATTGCTCCATA	13006
Qy	27	LysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyValaGlySerLysProGluLeu	46
Db	13007	AAATTTGCCCTTGGAGCAAAAGTTACAGCAGGCTTGCTCTGGGGCAAAAGCTGACAGTT	13066

QY 47 GlyLysArgSerAlaGlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsn 66
 DB 13067 GGACAGCTGCACAGACAGACATGATGTTAATCAAAATCACTTCTGCGCAT 13126
 QY 67 MetLeuPheIleThrGlyGlyMetGlyGlyLysThrCysThrGlyAlaAlaProVal 86
 DB 13127 ATGCTGTTATTTACTGCTGATGGGGGCGGACAGAAACCGGACCTCCAGTTATT 13186
 QY 87 AlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValSerThrProPheArg 106
 DB 13187 GCCCAAGTTCAGAAAGAGCTGATTTTAAACAGTGGAGTTGACAAAGCTTTTAAAG 13246
 QY 107 SerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValGlyGluLeuAlaLysTyr 126
 DB 13247 TTGGAAGCTCTTAAGAGTGAAGCTGCTGAGCAGGAATTAATTAAGAAAGTCT 13306
 QY 127 ValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaAspLysSerThr 146
 DB 13307 GTAGATACATGATGATTAATCAAAATCAAAAGCTTTTAACTGTTGTTGACAAAGAAC 13366
 QY 147 ThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyVal 166
 DB 13367 ACCATTAAAGATGCTTTTAAAGCGGACAGATGATGTTCTTAGAATGGCGCTTCAAGTATT 13426
 QY 167 ThrAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
 DB 13427 GCAGGCTTATTATTGAGCATGAGAGCTTAATATT 13462

RESULT 7

US-09-830-228-3

Sequence 3, Application US/09830228

GENERAL INFORMATION:

APPLICANT: Gil Choi et. al.

TITLE OF INVENTION: Borrelia burgdorferi Polynucleotides and Sequences

NUMBER OF SEQUENCES: 155

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/830,228

FILING DATE: 24-Apr-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US98/12764

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB370PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 11309 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-830-228-3

Alignment Scores:

Pred. No.: 4 92e-41 Length: 111309
 Score: 506.00 Matches: 101

Percent Similarity: 74.42% Conservative: 27
 Best Local Similarity: 58.72% Mismatches: 44
 Query Match: 57.18% Indels: 0
 DB: 32 Gaps: 0

US-09-770-509-2 (1-178) x US-09-830-228-3 (1-111309)

QY 7 ValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaProHis 26
 DB 12947 GTTGATTTATTGCTGCTAATACCATCTTCAGGCTCTCCAACTTATTGCTCCCTA 13006
 QY 27 LysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeu 46
 DB 13007 AAAATGCCCTTGAGCAAAAGTTACAGCAGGCGCTGCTGCGGGGAAAGCCGAGATT 13066
 QY 47 GlyLysArgSerAlaGlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsn 66
 DB 13067 GGACAGCTGCACAGACAGACATGATGTTAATCAAAATCACTTCTGCGCAT 13126
 QY 67 MetLeuPheIleThrGlyGlyMetGlyGlyLysThrCysThrGlyAlaAlaProVal 86
 DB 13127 ATGCTGTTATTTACTGCTGATGGGGGCGGACAGAAACCGGACGCTCCAGTATT 13186
 QY 87 AlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValSerThrProPheArg 106
 DB 13187 GCCCAAGTTCAGAAAGAGCTGATTTTAAACAGTGGAGTTGACAAAGCTTTTAAAG 13246
 QY 107 SerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGlyLeuAlaLysTyr 126
 DB 13247 TTGGAAGCTCTTAAGAGTGAAGCTGCTGAGCAGGAATTAATTAAGAAAGTCT 13306
 QY 127 ValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaAspLysSerThr 146
 DB 13307 GTAGATACATGATGATTAATCAAAATCAAAAGCTTTTAACTGTTGTTGACAAAGAAC 13366
 QY 147 ThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyVal 166
 DB 13367 ACCATTAAAGATGCTTTTAAAGCGGACAGATGATGTTCTTAGAATGGCGCTTCAAGTATT 13426
 QY 167 ThrAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
 DB 13427 GCAGGCTTATTATTGAGCATGAGAGCTTAATATT 13462

RESULT 8

PCT-US98-12764-1

Sequence 1, Application PC/TUS9812764

GENERAL INFORMATION:

APPLICANT: Gil Choi et. al.

TITLE OF INVENTION: Borrelia burgdorferi Polynucleotides and Sequences

NUMBER OF SEQUENCES: 155

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US98/12764

FILING DATE: Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB370PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 910715 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 PCR-US98-12764-1

Alignment Scores:

Pred. No.: 6,2e-40 Length: 910715
 Score: 506.00 Matches: 101
 Percent Similarity: 74.42% Conservative: 27
 Best Local Similarity: 58.72% Mismatches: 44
 Query Match: 57.18% Indels: 0
 Gaps: 0

US-09-770-509-2 (1-178) x PCT-US98-12764-1 (1-910715)

QY 7 ValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaProHis 26
 Db 602333 GTTGAATTATTGCTGCTATACCGATCTTCAGGCTCCCAACTCTATTCTCCCATTA 602392
 QY 27 LysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeu 46
 Db 602393 AAAATTGCCCTTGAGCAAAAGTTACAGCAGGCTTGCTGCGGCAAGCCTGAGATT 602452
 QY 47 GlyLysArgSerAlaGlnGlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsn 66
 Db 602453 GGACAAAGCTGCAGCAAGCAAGCATGATTATACGAAATCATCTTGCTGCTCCGAT 602512
 QY 67 MetLeuPheIleThrGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProValVal 86
 Db 602513 ATGCTGTTATTACTGCTGCTGATGGGGCGGAGCAAGAACCGGAGCCTCCCATTTATT 602572
 QY 87 AlaSerValAlaArgGluLeuGlyLysLeuThrValGlyValSerThrProPheArg 106
 Db 602573 GCGCAAGTTCGCAAAAGAGCTTGATTTTAAAGTTGAGTTGTAACAAAGCCTTTTAAAG 602632
 QY 107 SerGlnIleProAsnArgThrArgLeuAlaAsnAlaGlyLysGluLeuAlaLysThr 126
 Db 602633 TTTGAAGGCTCTTAAGAGATTGACACTTCTGAGCAGGGAATTAATTAAGTAAGAGTCT 602692
 QY 127 ValAspThrLeuIleValProAsnGlnAsnLeuAlaLeuAlaAspLysSerThr 146
 Db 602693 GTAGATACATTTGATTTATTCCTCAAAATCAAAAGCCTTTACTGTTGACAAAAGAAC 602752
 QY 147 ThrMetLeuGlnAlaPheArgThrArgLeuAlaAspAspValLeuLeuGlyValLysGlyVal 166
 Db 602753 ACCATTAAAGATGCTTTTAAAGCTGCAGATGATGTTCTTAGAATGGCGTTCAAGGTATT 602812
 QY 167 ThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
 Db 602813 GCAGGCTTATTATTGACATGAGAGGTTAATATT 602848

RESULT 9

US-09-830-228-1

Sequence 1, Application US/09830228

GENERAL INFORMATION:

APPLICANT: Gili Choi et. al.

TITLE OF INVENTION: Borrelia burgdorferi Polynucleotides and Sequences

NUMBER OF SEQUENCES: 155

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/830, 228

FILING DATE: 24-Apr-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US98/12764

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB370PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 910715 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-830-228-1

Alignment Scores:

Pred. No.: 6,2e-40 Length: 910715
 Score: 506.00 Matches: 101
 Percent Similarity: 74.42% Conservative: 27
 Best Local Similarity: 58.72% Mismatches: 44
 Query Match: 57.18% Indels: 0
 Gaps: 0

US-09-770-509-2 (1-178) x US-09-830-228-1 (1-910715)

QY 7 ValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaProHis 26
 Db 602333 GTTGAATTATTGCTGCTATACCGATCTTCAGGCTCCCAACTCTATTCTCCCATTA 602392
 QY 27 LysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeu 46
 Db 602393 AAAATTGCCCTTGAGCAAAAGTTACAGCAGGCTTGCTGCGGCAAGCCTGAGATT 602452
 QY 47 GlyLysArgSerAlaGlnGlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsn 66
 Db 602453 GGACAAAGCTGCAGCAAGCAAGCATGATTATACGAAATCATCTTGCTGCTCCGAT 602512
 QY 67 MetLeuPheIleThrGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProValVal 86
 Db 602513 ATGCTGTTATTACTGCTGCTGATGGGGCGGAGCAAGAACCGGAGCCTCCCATTTATT 602572
 QY 87 AlaSerValAlaArgGluLeuGlyLysLeuThrValGlyValSerThrProPheArg 106
 Db 602573 GCGCAAGTTCGCAAAAGAGCTTGATTTTAAAGTTGAGTTGTAACAAAGCCTTTTAAAG 602632
 QY 107 SerGlnIleProAsnArgThrArgLeuAlaAsnAlaGlyLysGluLeuAlaLysThr 126
 Db 602633 TTTGAAGGCTCTTAAGAGATTGACACTTCTGAGCAGGGAATTAATTAAGTAAGAGTCT 602692
 QY 127 ValAspThrLeuIleValProAsnGlnAsnLeuAlaLeuAlaAspLysSerThr 146
 Db 602693 GTAGATACATTTGATTTATTCCTCAAAATCAAAAGCCTTTACTGTTGACAAAAGAAC 602752
 QY 147 ThrMetLeuGlnAlaPheArgThrArgLeuAlaAspAspValLeuLeuGlyValLysGlyVal 166
 Db 602753 ACCATTAAAGATGCTTTTAAAGCTGCAGATGATGTTCTTAGAATGGCGTTCAAGGTATT 602812
 QY 167 ThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
 Db 602813 GCAGGCTTATTATTGACATGAGAGGTTAATATT 602848

RESULT 10

US-09-663-779-1189/c
 ; Sequence 1189, Application US/09663779
 ; GENERAL INFORMATION:
 ; APPLICANT: Corbin, David R.
 ; APPLICANT: Melver, Thomas M.
 ; APPLICANT: Shukla, Hridayabhajan
 ; TITLE OF INVENTION: BACILLUS THURINGIENSIS CHROMOSOMAL GENOME SEQUENCES AND USE
 ; FILE REFERENCE: 38-21(51376)B
 ; CURRENT APPLICATION NUMBER: US/09/663,779
 ; CURRENT FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: 1999-09-17
 ; PRIOR FILING DATE: 60/154,678
 ; NUMBER OF SEQ ID NOS: 8283
 ; SEQ ID NO 1189
 ; LENGTH: 7375
 ; TYPE: DNA
 ; ORGANISM: Bacillus thuringiensis
 ; US-09-663-779-1189

Alignment Scores:
 Pred. No.: 2,38e-42 Length: 7375
 Score: 505.00 Matches: 100
 Percent Similarity: 74.29% Conservative: 30
 Best Local Similarity: 57.14% Mismatches: 45
 Query Match: 57.06% Indels: 0
 DB: 26 Gaps: 0

US-09-770-509-2 (1-178) x US-09-663-779-1189 (1-7375)

QY 4 LeuGlulGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
 DB 1469 GTACAAAGCTAGACTTATCGCTGTAACACTGATGCAACAGCATTAATCATCAAAA 1410
 QY 24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
 DB 1409 GCTGAACAAACAAATGCAATGCGTGAACCGGAGCTTGTGTCAGCGGCAAC 1350
 QY 44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
 DB 1349 CCGTAAGTAAAGGAAAGCTGCGACAGAAAGTAAAGAACAGATCCAGAGCATTCGT 1290
 QY 64 AspSerAsnMetLeuPheIleThrGlyLysMetGlyGlyLysThrCysThrGlyAlaAla 83
 DB 1289 GGTGCGGATATGCTTCTGTAATCGCGGTATGGCGGTGGAAGTGAAGTGTGACGT 1230
 QY 84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValAlaSerThr 103
 DB 1229 CCAAGTGTCTCAAGTGAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1170
 QY 104 ProPheArgSerGlyLysProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
 DB 1169 CCAATTTCTTTTGAAGGACGTAAAGGTCGACGACGACGACGATCTGTATTCAGCATTT 1110
 QY 124 AlaLysThrValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaAsp 143
 DB 1109 AAAGAAATGTAGTACTCTTATTTGTAATTCACAAACGATCCGTTATTAAGATTTGTTGAT 1050
 QY 144 LysSerThrThrMetLeuGluAlaPheArgThrValAlaAspValLeuGluGlyVal 163
 DB 1049 AAAAATACGCCAATGTAAGAGCATTCGCTGAAGCTGATTAACGTTATTAAGCTCAAGGTT 990
 QY 164 LysGlyValThrAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
 DB 989 CAAGGTATTCGATTTAATTCGAACGCCAGGTTTAATTAATTAATTA 945

RESULT 11
 US-09-754-468-42
 ; Sequence 42, Application US/09754468
 ; GENERAL INFORMATION:
 ; APPLICANT: Iversen, Patrick L.
 ; TITLE OF INVENTION: Antisense Antibacterial Cell Division
 ; TITLE OF INVENTION: Composition and Method

; FILE REFERENCE: 0450-0033.30
 ; CURRENT APPLICATION NUMBER: US/09/754,468
 ; CURRENT FILING DATE: 2001-01-04
 ; PRIOR APPLICATION NUMBER: US 60/174,484
 ; PRIOR FILING DATE: 2000-01-04
 ; NUMBER OF SEQ ID NOS: 112
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 42
 ; LENGTH: 1864
 ; TYPE: DNA
 ; ORGANISM: Bartonella henselae
 ; US-09-754-468-42

Alignment Scores:
 Pred. No.: 7.4e-43 Length: 1864
 Score: 503.00 Matches: 103
 Percent Similarity: 73.71% Conservative: 26
 Best Local Similarity: 58.86% Mismatches: 46
 Query Match: 56.84% Indels: 0
 DB: 29 Gaps: 0

US-09-770-509-2 (1-178) x US-09-754-468-42 (1-1864)

QY 4 LeuGlulGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
 DB 223 CTTCAAGAGTGTGATTTGTTGCTTAATACAGATCCGACGCTTGTGCTATGTCACAAAG 282
 QY 24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
 DB 283 GCTGAACGTTATCAAGCTTGTGTCAGCGGTTACACAAAGTTTAAAGTGTGCTGCTTTA 342
 QY 44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
 DB 343 CCGGAAGTGGACACAGCGCTGACAGAGATGATTTATGATTAATTAATTAATTAATTAATTA 402
 QY 64 AspSerAsnMetLeuPheIleThrGlyLysMetGlyGlyLysThrCysThrGlyAlaAla 83
 DB 403 GATTCCTCATATGATTTCTTACTGCTGATGAGAGAGAGTACGGGAACAGAGACACA 462
 QY 84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValAlaSerThr 103
 DB 463 CCGTGTGTGACAGCTGCTCTGCTGTAAGAAAGTAAATTTTGAACCGTGTGTGTGTAACAG 522
 QY 104 ProPheArgSerGlyLysProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
 DB 523 CCAATTTCAATTTGAAGGCGCTGCGCTATGTAAGAAAGCGACAGGCTGTGATTAAGCATTA 582
 QY 124 AlaLysThrValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaAsp 143
 DB 583 CAAAATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 642
 QY 144 LysSerThrThrMetLeuGluAlaPheArgThrValAlaAspValLeuGluGlyVal 163
 DB 643 GAAAAACACACCTTTGCTGATGCTTTGCTATGCTGATGCTGATGCTGATGCTGATGCTG 702
 QY 164 LysGlyValThrAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
 DB 703 GCTTCATTAACAGATCATGATTAAGAAGGCTCATTAACCTT 747

RESULT 12
 US-09-974-300-180
 ; Sequence 180, Application US/09974300
 ; GENERAL INFORMATION:
 ; APPLICANT: Berka, Randy M.
 ; APPLICANT: Clausen, Ib Groth
 ; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
 ; FILE OF INVENTION: Expression
 ; FILE REFERENCE: 10085.500-05
 ; CURRENT APPLICATION NUMBER: US/09/974,300
 ; CURRENT FILING DATE: 2001-10-05
 ; PRIOR APPLICATION NUMBER: 09/680,598
 ; PRIOR FILING DATE: 2000-10-06
 ; PRIOR APPLICATION NUMBER: 60/279,526

PRIOR FILING DATE: 2001-03-27
 NUMBER OF SEQ ID NOS: 8481
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 180
 LENGTH: 1134
 TYPE: DNA
 ORGANISM: Bacillus licheniformis
 US-09-974-300-180

Alignment Scores:
 Pred. No.: 5.19e-43 Length: 1134
 Score: 502.00 Matches: 97
 Percent Similarity: 74.01% Conservative: 34
 Best Local Similarity: 54.80% Mismatches: 46
 Query Match: 56.72% Indels: 0
 Gaps: 0

US-09-770-509-2 (1-178) x US-09-974-300-180 (1-1134)

QY 2 SerGlnLeuGlnGlyValAlaGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
 DB 97 AATGAGGTTTCAGGAGTCGAGTTTATGCAAGTCAACGAGATGCTCAAGCTCTCAACCTG 156
 QY 22 SerLeuAlaProHisLysIleThrLeuGlnGlyLysAspIleThrLysGlyLeuGlnGlyArg 41
 DB 157 TCAAAAGCCGGAACGAAATGCGAGATGCTGCGAAGCTGACGCGGCGGCTGCGCCCGGA 216
 QY 42 SerLysProGlnLeuGlnGlyLysArgSerAlaGlnGlnGlnLysValAspIleGlnArgMet 61
 DB 217 GCCAATCCGGAAGTGGGCAAAAAGCCGAGAGCAAGCAAAAGCAAAATGTAAGAACGCG 276
 QY 62 LeuGlnAspSerAsnMetLeuPheIleThrGlyLysMetGlyGlyGlyThrCysThrGly 81
 DB 277 CTGAAGAGTGTCATGATGTTGCTGTCACAGCCGGAATGGCGGCGGAGCAAGAGCGGCG 336
 QY 82 AlaAlaProValAlaAsnValAlaArgGlnLeuGlnGlyIleLeuThrValGlyValAla 101
 DB 337 GGGGCACTGTATCTCCACCAATCGCAAAAGATGTGGCGCATTCGCTGCGCGCTGTC 396
 QY 102 SerThrProPheArgSerGlnGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
 DB 397 ACAAGAGCCGTTTACCTTGAAGAGCAAAAGACAGCTTCAGGCTGCAAGCGGATTTTCA 456
 QY 122 GluLeuAlaLysThrValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeu 141
 DB 457 GCAATGAGAGAGCCGTCGACACCTGATGCTATTCGCAATGACCGCTTCTTGAATTC 516
 QY 142 AlaAspLysSerThrThrMetLeuGlnAlaPheArgThrAlaAspAspValLeuLeuGlu 161
 DB 517 GTGCAATAAACACACACCGATGCTTGAAGCTTCGTAAGCGGCAACGCTTCTCCGCCAA 576
 QY 162 GlyValLysGlyValThrAspLeuIleValAlaArgProGlnLysLeuIleAsnLeu 178
 DB 577 GGTGTTCAGGCGATTCACACGATGCGCAACGCTGACTGATCAACCTT 627

RESULT 13

US-60-045-649-854
 Sequence 854, Application US/60045649
 GENERAL INFORMATION:

APPLICANT: Lagace, Robert E.
 APPLICANT: Corely, Neil C.
 APPLICANT: Russo, Frank D.
 APPLICANT: Hann, Amy L.
 TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
 NUMBER OF SEQUENCES: 1466
 CORRESPONDENCE ADDRESS:
 STREET: 3174 PORTER DRIVE
 CITY: PALO ALTO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/60/045,649
 FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
 NAME: CERRONE, MICHAEL C.
 REGISTRATION NUMBER: 39,132
 REFERENCE/DOCKET NUMBER: PM-0003P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 855-0555
 TELEFAX: (415) 845-4166
 INFORMATION FOR SEQ ID NO: 854:

SEQUENCE CHARACTERISTICS:

LENGTH: 3993 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 IMMEDIATE SOURCE:
 CLONE: EF10854
 US-60-045-649-854

Alignment Scores:

Pred. No.: 1.03e-41 Length: 3993
 Score: 496.00 Matches: 100
 Percent Similarity: 73.14% Conservative: 28
 Best Local Similarity: 57.14% Mismatches: 47
 Query Match: 56.05% Indels: 0
 Gaps: 0

US-09-770-509-2 (1-178) x US-60-045-649-854 (1-3993)

QY 4 LeuGlnGlyValAlaGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
 DB 2707 GTTAAAGCGCGGATTTATATACACAGCAATACAGAGCTTCAGCAAGCATTAACATCAAAA 2766
 QY 24 AlaProHisLysIleThrLeuGlnGlyLysAspIleThrLysGlyLeuGlnGlyArgSerLys 43
 DB 2767 GCAGAAACAGTATTCATTAAGCCCTTAATACACTCGGTGTTAGTCCGCTTACAA 2826
 QY 44 ProGlnLeuGlnGlyLysArgSerAlaGlnGlnGlnLysValAspIleGlnArgMetLeuGln 63
 DB 2827 CTTGAAGTTGGCCAAAAGCTGCGAGAGAAAGTGAACAAGTATTCACAAATCATTTACAA 2886
 QY 64 AspSerAsnMetLeuPheIleThrGlyLysMetGlyGlyGlyThrCysThrGlyAlaAla 83
 DB 2887 GGGCGGATATGATTTATTCATTAAGCTGCTGATGCTGCGGAGAACTGTACAGGTCTGCG 2946
 QY 84 ProValAlaAlaSerValAlaArgGlnLeuGlnGlyIleLeuThrValGlyValAlaSerThr 103
 DB 2947 CCAGTAGTTCAAAATCGCTTAAAGAAATTAAGCGCTTTTAAACGTTGCTGTAATCACTCGT 3006
 QY 104 ProPheArgSerGlnGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGlnLeu 123
 DB 3007 CCATTATGTTTGAAGGTCCAAACGTCGTGCTTTGCGCGCGAAGAAATTCCTTATTA 3066
 QY 124 AlaLysThrValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaAsp 143
 DB 3067 AAGAAACAGTGTATACATTAATATATCAACCAACCGCTTATTTGAAGTCTGTAC 3126
 QY 144 LysSerThrThrMetLeuGlnAlaPheArgThrAlaAspAspValLeuLeuGlnGlyVal 163
 DB 3127 AAGAAACAGCAATGCTTGAAGCATTTAAGAACTGATTAATGTTATGTTACGTTCAAGTGT 3186
 QY 164 LysGlyValThrAspLeuIleValAlaArgProGlnLysLeuIleAsnLeu 178
 DB 3187 CAAGGATTTCAATTAATCACTGACACAGGATTAACGTAACCTTA 3231

RESULT 14
US-60-046-653-853
; Sequence 853, Application US/60046653
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert E.
; APPLICANT: Corley, Neil C.
; APPLICANT: Russo, Frank D.
; APPLICANT: Hann, Amy L.
; APPLICANT: Heath, Joe D.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
; NUMBER OF SEQUENCES: 149
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/046.653
; FILING DATE: HERewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0003-1 P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 853:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3993 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: EFALC853
; US-60-046-653-853
Alignment Scores:
Pred. No.: 1.03e-41 Length: 3993
Score: 496.00 Matches: 100
Percent Similarity: 73.14% Conservative: 28
Best Local Similarity: 57.14% Mismatches: 47
Query Match: 56.05% Indels: 0
Gaps: 0
US-09-770-509-2 (1-178) x US-60-046-653-853 (1-3993)
QY 4 LeuGlUgLYVaLgLuPheILeVaLAlAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
Db 2707 GTTAAAGCGGTGGAATTTATACACGCAATACAGCTTCAAGCATTAACATTCAAAA 2766
QY 24 AlaProHIsLysILeHrLeuGlyLysAspILeHrLysGlyLeuGlyAlaGlySerLys 43
Db 2767 GCAGAAACAGTGAATTCATTTAGCCCTAAATACACTCGGTGTAGTGGCGGTTCACA 2826
QY 44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspILeGlnArgMetLeuGln 63
Db 2827 CCGTGAAGTTGGCCAAAAGCTGCAGAAACAGTGAACAGTATTCAGATCATTCACAA 2886
QY 64 AspSerAsnMetLeuPheILeHrLysGlyLysGlyLysGlyLysGlyLysGlyLysGly 83
Db 2887 GCGCGGATATGATTTTCATTACTGCTGATGGTGGTGGGAGACTGCTACAGGTGCTGCG 2946
QY 84 ProValAlaIAserValAlaArgGluLeuGlyIleLeuThrValGlyValSerThr 103

RESULT 15
PCT-US02-03987-6542
; Sequence 6542, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elitra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for identifying the target of a compound which inhibits
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA 0288VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; CURRENT FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FASTSeq for Windows Version 4.0
; SEQ ID NO: 6542
; LENGTH: 1233
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1233)
; PCT-US02-03987-6542
Alignment Scores:
Pred. No.: 3.19e-42 Length: 1233
Score: 495.00 Matches: 100
Percent Similarity: 73.14% Conservative: 28
Best Local Similarity: 57.14% Mismatches: 47
Query Match: 55.93% Indels: 0
Gaps: 0
US-09-770-509-2 (1-178) x PCT-US02-03987-6542 (1-1233)
QY 4 LeuGlUgLYVaLgLuPheILeVaLAlAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
Db 106 GTTAAAGCGGTGGAATTTATACACGCAATACAGCTTCAAGCATTAACATTCAAAA 165
QY 24 AlaProHIsLysILeHrLeuGlyLysAspILeHrLysGlyLeuGlyAlaGlySerLys 43
Db 166 GCAGAAACAGTGAATTCATTTAGCCCTAAATACACTCGGTGTAGTGGCGGTTCACA 225
QY 44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspILeGlnArgMetLeuGln 63
Db 226 CCGTGAAGTTGGCCAAAAGCTGCAGAAACAGTGAACAGTATTCAGATCATTCACAA 285
QY 64 AspSerAsnMetLeuPheILeHrLysGlyLysGlyLysGlyLysGlyLysGlyLysGly 83
Db 286 GCGCGGATATGATTTTCATTACTGCTGATGGTGGTGGGAGACTGCTACAGGTGCTGCG 345
QY 84 ProValAlaIAserValAlaArgGluLeuGlyIleLeuThrValGlyValSerThr 103
Db 346 CCAGTAGTTGCAAAAATCGCTTAAGAGAGTGGCGCTTAACAGTTGGTGTAGTAACTGCT 405
QY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValSerThr 123

```
Db      406  CATTAGTTTGAAGTCCAAAACGTGCTGTTTCCGCTGAAGGAATTCCTTATTA 465
OY      124  ALALSTYrValAspThrLeuIleValProAsnGlnAsnLeuAlaLeuAlaasp 143
Db      466  AAGAAACGCTGATACATTAATATCTCAACACCGCTTATTGAAAGTCGTTGAC 525
OY      144  LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
Db      526  AAGAAACGCGCAATGCTTGAAGCATTTAGAGAGCTGATATATGTTAGTCACAGGTGTT 585
OY      164  LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
Db      586  CAAGGATTTCAGATTAATCACTGACACGATTACGTAACCTTA 630
```

Search completed: June 2, 2003, 11:06:15
Job time : 3441 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - nucleic search, using frame_plus_p2n model

Run on: June 2, 2003, 09:41:39 ; Search time 740 Seconds
(without alignments)
1503.471 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDQALG.....LLEGVKGVTDLIVRGLINTL 178

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7080287 seqs, 3125191874 residues

Total number of hits satisfying chosen parameters: 14160574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=x1h
-CG/cgcn2_1/USPTO.spool/US09770509/runat_27052003_083652_18961/app.query.fasta_1.327
-DB=Pending_Patents_NA_New -QFMT=fastap -SUFFIX=trnpn -MINMATCH=0.1 -LOOPEC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09770509 @CGCN1.1.671 @runat_27052003_083652_18961 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -MATT=DSPBLOCK-100 -LONCLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_New.*

1: /cgcn2_6/pdata/1/pna/US06_NEW.COMB.seq.*
2: /cgcn2_6/pdata/1/pna/US06_NEW.COMB.seq.*
3: /cgcn2_6/pdata/1/pna/US07_NEW.COMB.seq.*
4: /cgcn2_6/pdata/1/pna/US08_NEW.COMB.seq.*
5: /cgcn2_6/pdata/1/pna/US09_NEW.COMB.seq.*
6: /cgcn2_6/pdata/1/pna/US09_NEW.COMB.seq.*
7: /cgcn2_6/pdata/1/pna/US09_NEW.COMB.seq.*
8: /cgcn2_6/pdata/1/pna/US10_NEW.COMB.seq.*
9: /cgcn2_6/pdata/1/pna/US10_NEW.COMB.seq.*
10: /cgcn2_6/pdata/1/pna/US60_NEW.COMB.seq.*
11: /cgcn2_6/pdata/1/pna/US60_NEW.COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	510	57.6	1176	9	US-10-282-122A-24440
2	510	57.6	213251	8	US-10-398-221-6
3	510	57.3	3011208	8	US-10-398-221-2058
4	507	57.3	1122	9	US-10-282-122A-15434
5	506	57.2	1215	9	US-10-282-122A-10897
6	505	57.1	1158	9	US-10-282-122A-10662
7	505	57.1	1170	9	US-10-282-122A-9479

8	505	57.1	1182	9	US-10-282-122A-17147	Sequence 17147, A
9	495	55.9	1230	6	US-09-134-000C-6285	Sequence 6285, Ap
10	495	55.9	1308	6	US-09-134-000C-1253	Sequence 1253, Ap
11	495	55.9	1308	6	US-09-134-000C-1253	Sequence 1253, Ap
12	494	55.8	1245	8	US-10-417-884-3641	Sequence 3641, Ap
13	494	55.8	1245	9	US-10-282-122A-21408	Sequence 21408, A
14	492	55.6	1161	9	US-10-282-122A-25784	Sequence 25784, A
15	490	55.4	1107	9	US-10-282-122A-16538	Sequence 16538, A
16	490	55.4	1260	1	PCT-US02-36122-83	Sequence 83, Appl
17	490	55.4	1260	1	PCT-US02-36123-5201	Sequence 5201, Ap
18	490	55.4	1754382	1	PCT-US02-36123-6651	Sequence 6651, Ap
19	486	54.9	1188	8	US-10-156-761-6101	Sequence 6101, Ap
20	486	54.9	1233	9	US-10-282-122A-17831	Sequence 17831, A
21	485	54.9	9025608	8	US-10-156-761-1	Sequence 1, Appl
22	485	54.8	1137	9	US-10-282-122A-26150	Sequence 26150, A
23	485	54.8	1140	9	US-10-282-122A-28464	Sequence 28464, A
24	482	54.5	1179	9	US-10-282-122A-35322	Sequence 35322, A
25	481	54.4	1140	9	US-10-282-122A-27607	Sequence 27607, A
26	480	54.2	1173	6	US-09-950-084-1797	Sequence 1797, Ap
27	480	54.2	1173	9	US-10-282-122A-7824	Sequence 7824, Ap
28	480	54.2	1182	9	US-10-282-122A-34548	Sequence 34548, A
29	480	54.2	1185	9	US-10-092-411A-1987	Sequence 1987, Ap
30	480	54.2	2893	10	US-60-009-861-190	Sequence 190, App
31	480	54.2	3931	9	US-10-329-624-342	Sequence 342, App
32	480	54.2	98958	6	US-09-950-084-7442	Sequence 7442, Ap
33	472	53.3	1260	9	US-10-282-122A-37929	Sequence 37929, A
34	471	53.2	1194	9	US-10-282-122A-31723	Sequence 31723, A
35	468	52.9	1185	9	US-10-282-122A-33518	Sequence 33518, A
36	466	52.7	29112	1	PCT-US02-3727-104	Sequence 104, App
37	466	52.7	29112	6	US-09-978-825-104	Sequence 104, App
38	466	52.7	29112	9	US-10-057-498-104	Sequence 104, App
39	465	52.5	1257	9	US-10-282-122A-40313	Sequence 40313, A
40	464	52.4	1185	9	US-10-282-122A-30463	Sequence 30463, A
41	464	52.4	1203	8	US-10-419-128-7557	Sequence 7557, Ap
42	464	52.4	1203	9	US-10-366-683-7557	Sequence 7557, Ap
43	464	52.4	1314	8	US-10-419-128-7784	Sequence 7784, Ap
44	464	52.4	1314	9	US-10-366-683-7784	Sequence 7784, Ap
45	464	52.4	1320	9	US-10-282-122A-38466	Sequence 38466, A

ALIGNMENTS

RESULT 1
US-10-282-122A-24440
: Sequence 24440, Application US/10282122A
: GENERAL INFORMATION:
: APPLICANT: Wang, Liangsu
: APPLICANT: Zamudio, Carlos
: APPLICANT: Malone, Cheryl
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl
: APPLICANT: Zyskind, Judith
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John
: APPLICANT: Carr, Grant
: APPLICANT: Yamamoto, Robert
: APPLICANT: Forsyth, R.
: TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
: FILE REFERENCE: ELITRA.034A
: CURRENT FILING DATE: 2003-02-20
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/230,335
: PRIOR FILING DATE: 2000-09-06
: PRIOR APPLICATION NUMBER: 60/230,347
: PRIOR FILING DATE: 2000-09-09
: PRIOR APPLICATION NUMBER: 60/242,578

```

: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/267,636
: PRIOR FILING DATE: 2001-02-09
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 78614
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 24440
: LENGTH: 1176
: TYPE: DNA
: ORGANISM: Listeria monocytogenes
US-10-282-122A-24440

Alignment Scores:
Pred. No.: 2,3e-48 Length: 1176
Score: 510.00 Matches: 98
Percent Similarity: 76.00% Conservative: 35
Best Local Similarity: 56.00% Mismatches: 42
Query Match: 57.63% Indels: 0
DB: Gaps: 0

US-09-770-509-2 (1-178) x US-10-282-122A-24440 (1-1176)
QY 4 LeuglugglYValIGluPheIleValAlaAsnThrAspCysGlnAlaIleuGlyArgSerLeu 23
Db 103 GTTCAGAGAGTGAATTTATCTCCGTTAATACAGACGCTCAACACTTAACCTAAGCAAAA 162
QY 24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
Db 163 GCAGAACCAAAATTCAAATCGGTACAAATTAACCGCGTGTAGTGGCGGAGCGTGA 222
QY 44 ProGluLeuGlyLysArgSerIleGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
Db 223 CCTGAATTTGGTAAATAAGCTCAGACAGAAAGCCGGAAACAAATTCAGAAAGCTTAAAA 282
QY 64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyIleThrCysThrGlyAlaAla 83
Db 283 GGCTCTGATATAGTGTATGCTAGTACTGCTGGATATGGCGGGGGAACCTGAGACTGGGGCTGCA 342
QY 84 ProValValAlaIleSerValAlaIleArgGluLeuGlyIleLeuThrValGlyValIleSerThr 103
Db 343 CCTGTATATGCTCAATTCGCAAAATAAAGGGCGCTTTAAACGTAGCTGTGTACTCGCA 402
QY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
Db 403 CCATTTGGTTTTGAAGACCAAAAGCTACGAAACAAGCCTTAACCTGGAACAGAACATG 462
QY 124 AlaIleStyValAspThrLeuIleValValProAsnGlnAsnLeuAlaLeuAlaAsp 143
Db 463 AAAAAGCGGGGAGTACCTTAATTTCTTAATGACCGTTTACTCAAAATTTGGAT 522
QY 144 LysSerThrTrpMetLeuGlnAlaIlePheArgTyrAlaAspAspValLeuLeuGluGlyVal 153
Db 523 AAAAATACACCGATGCTTGAAGCTTTCCGTGAAGCAGATAAAGTTTACGTCAAGGTGA 582
QY 164 LysGlyValIleThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
Db 583 CAAGGATTTCTGATTTGATTCGCCCTCTCGTTTAATTAACCTTA 627

RESULT 2
US-10-398-221-6
: Sequence 6, Application US/10398221
: GENERAL INFORMATION:
: APPLICANT: KUNST, Frederik
: APPLICANT: GLASER, Philippe
: TITLE OF INVENTION: Listeria innocua, genome and applications
: FILE REFERENCE: 344 702 - US
: CURRENT APPLICATION NUMBER: US/10/398,221

```

```

CURRENT FILING DATE: 2003-03-27
PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: FR 00/12 697
PRIOR FILING DATE: 2000-10-04
NUMBER OF SEQ ID NOS: 4025
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 213251
TYPE: DNA
ORGANISM: Listeria innocua
US-10-398-221-6

Alignment Scores:
Pred. No.: 2 78e-45 Length: 213251
Score: 510.00 Matches: 98
Percent Similarity: 76.008 Conservative: 35
Best Local Similarity: 56.004 Mismatches: 42
Query Match: 57.63% Indels: 0
DB: 8 Gaps: 0

US-09-770-509-2 (1-178) x US-10-398-221-6 (1-213251)
QY 4 LeunglglYvalGlupheleValaAlaasntrhAspCysGlnAlaLeuglYargSerleu 23
Db 170157 GTTCAGAGGATGGAATTTATCTCCGTTATACAGAGCGCTCAAGCACTTAATTATGACAAA 170218
QY 24 AlaProhlsYlIaethrlenuglYlYsAspIlethrlYglYleuGlYlaGlYserlYs 43
Db 170217 GCAGAAACAAATTCACAAATCGGTACAAAATTAACCGGTGGTTAGCGCGGGTGTCTGA 170276
QY 44 ProgluLeuglYlYsargSerAlaGluglInglnlYsValasPlleGlnarYmetLeugln 63
Db 170277 CCTGAATTTGTAATAAAACACACAGACAGAAAGTCGCGACAAATTAAGACAGCTTTAAA 170336
QY 64 AspSerAsmetLeuphelethrlYglYglYmetGlYglYlYthrcYsthrGlYAlaIa 83
Db 170337 GCGCTTGATGTGATGTCTTAACCTGCTGGATGGCGGGGGAACGTGAACGTGTCTGA 170396
QY 84 ProValValAlaSerValAlaIaArggluLeuglYlIeLeuThValGlYValSerThr 103
Db 170397 CCTGTATCGGCTCAAAATCGCAAAAGAGGGCGGCTTAACGTGCGTGTGTATACAGA 170456
QY 104 ProPheArgSerGluglYlYProAsnArghrArGlueAlaAsnAlaGlYAlYlYsgluleu 123
Db 170457 CCATTTGGTTTGAAGGACCAAAAGCTGCAAAACAAAGCCCTAAGCTGAACAGACAGAAATG 170516
QY 124 AlaLysrYlValAspThrLeuIleValAlaProAsnGlnAsnLeuLeuAlaIaAsp 143
Db 170517 AAAGAGCGGTGAGTACTTAAATTGTATACCTTAACGACCGTTTACTTCAAAATGTGTGAT 170576
QY 144 LysSerThrTrhMetLeuGlulAlaIaPheArgrYrAlaAspAspValleuLeuglYlYAl 163
Db 170577 AAAAATAGCCAAATCTGTAAGCTTTCGCTGAGACAGATAAATGTTTACGTCAAGGGGTA 170636
QY 164 LysGlYValAlThrasPleulleValaIaArgProGlYleuIleAsnleu 178
Db 170637 CAAGGATTTCTGATTGATTGCGCTTCCTGTTTAATTAACCTTA 170681

RESULT 3
US-10-398-221-2058/c
Sequence 2058, Application US/10398221
GENERAL INFORMATION:
APPLICANT: KUNST, Frederik
APPLICANT: GLASER, Philippe
TITLE OF INVENTION: Listeria innocua, genome and applications
FILE REFERENCE: 344 702 - US
CURRENT APPLICATION NUMBER: US/10/398, 221
CURRENT FILING DATE: 2003-03-27
PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: FR 00/12 697
PRIOR FILING DATE: 2000-10-04

```

```
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2058
; LENGTH: 3011208
; TYPE: DNA
; ORGANISM: Listeria innocua
US-10-398-221-2058

Alignment Scores:
Pred. No.: 1.03e+43      Length: 3011208
Score: 510.00           Matches: 98
Percent Similarity: 76.00%      Conservative: 35
Best Local Similarity: 56.00%    Mismatches: 42
Query Match: 57.63%           Indels: 0
Gaps: 0

US-09-770-509-2 (1-178) x US-10-398-221-2058 (1-3011208)

QY 4 LeuGIuGIyValGIuPheIleValAlaAsnThrAspCysGlnAlaLeuGIyArgSerLeu 23
Db 2161845 GTTCAGAGAGAGATATTCCTCCGTATACAGACGCTCAACACCTTAATTAGCAAAA 2161786
QY 24 AlaProHisLysIleThrLeuGIyLysAspIleThrLysGIyLeuGIyAlaGIySerLys 43
Db 2161785 GCAGAAACAAATTAACAATCGGTCAAAATTAACCGCTTACGGCGGGGTGTGTA 2161726
QY 44 ProGIuLeuGIyLysArgSerAlaGIuGlnGlnLysValAspIleGlnArgMetLeuGln 63
Db 2161725 CCTGAATTTGGTAAAAAGACACACAGAAAGTCCGCAACAAATTGAGAAAGCTTTAAA 2161666
QY 64 AspSerAsnMetLeuPheIleThrGIyMetGIyLysThrCysThrGIyAlaAla 83
Db 2161665 GGCTCTGATATGATATGATCTGATCTGATGCGGCGGAGACTGGAAGTGGTGTGCA 2161606
QY 84 ProValAlaLysSerValAlaArgLysLeuGIyLysLeuThrValGIyValSerThr 103
Db 2161605 CCTGTATTCGTCAAAATGCAAAAGAAATGCGCTCACTGCTGTTGTTGTTACACA 2161546
QY 104 ProPheArgSerGIyLysProAsnArgThrArgLeuAlaAsnAlaGIyValLysGlnLeu 123
Db 2161545 CCATTGTGTTTGAAGAGCAACAAAGCTACGAAACCCCTAAGTGAACAGAAAGCAATG 2161486
QY 124 AlaLysThrValAlaPheThrLeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaAsp 143
Db 2161485 AAAGAAACGCGTGAATCTTAATGTTATCCCTAAGACGCTTACTTCAAAATGTTGAT 2161426
QY 144 LysSerThrThrMetLeuGIuAlaPheArgThrAlaAspAspValLeuLeuGIyVal 163
Db 2161425 AAAAATACGCCAAATGCTGAAGCTTCCGTGAAGCAGATAATGTTTACGTCAAGGGGTA 2161366
QY 164 LysGIyValThrAspLeuIleValAlaArgProGIyLeuIleAsnLeu 178
Db 2161365 CAAGGCAATTCATTTATTCCTGCTGTTTAATTACTTA 2161321

RESULT 4
US-10-282-122A-15434
; Sequence 15434, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Juddith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUTRA-034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
```

```
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15434
; LENGTH: 1122
; TYPE: DNA
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-15434

Alignment Scores:
Pred. No.: 4.77e+48      Length: 1122
Score: 507.00           Matches: 98
Percent Similarity: 76.00%      Conservative: 35
Best Local Similarity: 56.00%    Mismatches: 42
Query Match: 57.29%           Indels: 0
Gaps: 0

US-09-770-509-2 (1-178) x US-10-282-122A-15434 (1-1122)

QY 4 LeuGIuGIyValGIuPheIleValAlaAsnThrAspCysGlnAlaLeuGIyArgSerLeu 23
Db 103 CTTAAGAAAGTGAATTAATTCATTAACACTGATTAAGCGGCAATTCCTTCGCA 162
QY 24 AlaProHisLysIleThrLeuGIyLysAspIleThrLysGIyLeuGIyAlaGIySerLys 43
Db 163 GCATCAACAGATTAACAATAGGTGATTAAGCTTAAAGGCTTAAGGACGAGACTAAT 222
QY 44 ProGIuLeuGIyLysArgSerAlaGIuGlnGlnLysValAspIleGlnArgMetLeuGln 63
Db 223 CCTGAATATGCTGAAGAGCTGCTGAGAAAGTAAGATGAATATCTCAGGCTTTAAA 282
QY 64 AspSerAsnMetLeuPheIleThrGIyMetGIyLysThrCysThrGIyAlaAla 83
Db 283 GGTGCTGATATGATATTAATTAACCTGCTGATGCTGATGCTGAACCTGCTCAGCA 342
QY 84 ProValAlaLysSerValAlaArgGIuLeuGIyLysLeuThrValGIyValSerThr 103
Db 343 CCTGTGCTGCAACAAATGCTTAATCCATGGAAATTAACAGTGTGAGTTTACAAAG 402
QY 104 ProPheArgSerGIyLysProAsnArgThrArgLeuAlaAsnAlaGIyValLysGlnLeu 123
Db 403 CCTTCCCTTTTGAAGAGAAAGAGAAAGTCCATGCTGATGCTGAGTAATTAACACATA 462
QY 124 AlaLysThrValAlaPheThrLeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaAsp 143
Db 463 AAAGAAAGATGATCTTATTAATTCAAATGCAATGCAATGCTTGTATTTAGAT 522
QY 144 LysSerThrThrMetLeuGIuAlaPheArgThrAlaAspAspValLeuLeuGIyVal 163
Db 523 AAAAAGCAACTTAATTAAGACCTTTAAATCAGACAGTGTGTTTAAAGCAAGGTTT 582
QY 164 LysGIyValThrAspLeuIleValAlaArgProGIyLeuIleAsnLeu 178
```

```

Db      583 CAAGGATATACAGATTGATTAACCTAATCTCGTCTTGAACCTT 627
      RESULT 5
US-10-282-122A-10897
; Sequence 10897, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10897
; LENGTH: 1215
; TYPE: DNA
; ORGANISM: Borrelia burgdorferi
US-10-282-122A-10897

Alignment Scores:
Pred. No.:      6,92e-48      Length:      1215
Score:          506.00      Matches:      101
Percent Similarity: 74.42%      Conservative: 27
Best Local Similarity: 58.72%      Mismatches:  44
Query Match:    57.18%      Indels:      0
Db:              9      Gaps:      0

US-09-770-509-2 (1-178) x US-10-282-122A-10897 (1-1215)
QY      7 ValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaProHis 26
      |||||
Db      154 GTGATATTATTATGTTGGCTAATACCGATCTTCAGGCTCTCCAAACTTGTATGCTCCATA 213
      |||||

QY      27 LysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeu 46
      |||||
Db      214 AAAATGGCCCTTGGAGCAAAAGTTACAGCAGGCGTTGGTGGGGAAGCCTGAGATT 273
      |||||

QY      47 GlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsn 66
      |||||
Db      274 GGACAAAGCTGGAGCGAGAGAGACATAGATGTTATGTAATCAATCATCTTCTGAGCGGAT 333
      |||||

```

```

Oy      67 MetLeuPheLIleThrclyGlyMetGlyGlyGlyThrcCysThrclyAlaAlaProValVal 86
Db      334 ATGGTGCTTATTACTCTGCTGGTATGGGGGGGGGACAGCAAGAAACCGAGACGCTCCAGTTATT 3333
Oy      87 AlaSerValAlaArgGluLeuGlyLileuThrValGlyValSerThrProPheArg 106
Db      394 GCGCAAGTGTCAAAAGAGCTGGTGGTATTTTAACAGTTGGAGGTGTACAAAGACCTTTTAAAG 4533
Oy      107 SerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyr 126
Db      454 TTGTGAAGGCTCAAGAGAAGTTGAGACTTGCTGAGCAGGAGATTAATATCAAGAAAGTCT 5133
Oy      127 ValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaAlaLeuAlaAspLysSerThr 146
Db      514 GTAGATACATTGATCATCATTTATTCCAATCAAAAGCTTTTAACTGCTGTTGACAAAGAACC 5733
Oy      147 ThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyVal 166
Db      574 ACCATTAAAGACTCTTTTAAAGCGTGCAATGATGTTCTTGAATGGCGCTTCAAGGTATT 6333
Oy      167 ThrAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
Db      634 GCAGGCGTTATTATTGAGCATGAGAGAGTTAATTATT 669

RESULT 6
US-10-282-122A-10662
? Sequence 10662, Application US/10282122A
? GENERAL INFORMATION:
? APPLICANT: Wang, Liangsu
? APPLICANT: Zamudio, Carlos
? APPLICANT: Malone, Cheryl
? APPLICANT: Haselbeck, Robert
? APPLICANT: Ohlsen, Kari
? APPLICANT: Zyskind, Judith
? APPLICANT: Wall, Daniel
? APPLICANT: Trawick, John
? APPLICANT: Carr, Grant
? APPLICANT: Yamamoto, Robert
? APPLICANT: Forsyth, R.
? APPLICANT: Xu, H.
? TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
? FILE REFERENCE: ELITRA.034A
? CURRENT APPLICATION NUMBER: US/10/282,122A
? CURRENT FILING DATE: 2003-02-20
? PRIOR APPLICATION NUMBER: 60/191,078
? PRIOR FILING DATE: 2000-03-21
? PRIOR APPLICATION NUMBER: 60/206,848
? PRIOR FILING DATE: 2000-05-23
? PRIOR APPLICATION NUMBER: 60/207,727
? PRIOR FILING DATE: 2000-05-26
? PRIOR APPLICATION NUMBER: 60/230,335
? PRIOR FILING DATE: 2000-09-06
? PRIOR APPLICATION NUMBER: 60/230,347
? PRIOR FILING DATE: 2000-09-09
? PRIOR APPLICATION NUMBER: 60/242,578
? PRIOR FILING DATE: 2000-10-23
? PRIOR APPLICATION NUMBER: 60/253,625
? PRIOR FILING DATE: 2000-11-27
? PRIOR APPLICATION NUMBER: 60/257,931
? PRIOR FILING DATE: 2000-12-22
? PRIOR APPLICATION NUMBER: 60/267,636
? PRIOR FILING DATE: 2001-02-09
? PRIOR APPLICATION NUMBER: 60/269,308
? PRIOR FILING DATE: 2001-02-16
? Remaining Prior Application data removed - See File Wrapper or PALM.
? NUMBER OF SEQ ID NOS: 78614
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 10662
? LENGTH: 1158
? TYPE: DNA
? ORGANISM: Bacillus anthracis
? US-10-282-122A-10662

```

Alignment Scores:

Pred. No.:	8,45e-48	Length:	1158
Score:	505.00	Matches:	100
Percent Similarity:	74.29%	Conservative:	30
Best Local Similarity:	57.14%	Mismatches:	45
Query Match:	57.06%	Indels:	0
DB:	9	Gaps:	0

US-09-770-509-2 (1-178) x US-10-282-122A-10662 (1-1158)

```

OY 4 LeuGlUGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
    ::::::::::::::::::::|
DB 103 GTCAAGAGGTAGACTTTATCGCTGTGATCTGATGCAACAGCATTAATCATCAAAA 162
    ::::::::::::::::::::|
OY 24 AlaProHIsLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyLysSerLys 43
    ::::::::::::::::::::|
DB 163 GCTGAACAGAAATGCAAAATGCTGCGAATAAAGTAAGAACATTCAGACACTTCCT 222
    ::::::::::::::::::::|
OY 44 ProGluLeuGlyLysArgSerAlaGluGlnGlyValAspIleGlnArgMetLeuGln 63
    ::::::::::::::::::::|
DB 223 CCTGAAGTAGAGAAAGAAAGCTGCGAATAAAGTAAGAACATTCAGACACTTCCT 282
    ::::::::::::::::::::|
OY 64 AspSerAsnMetLeuPheIleThrLysGlyMetGlyGlyThrCysThrGlyAlaAla 83.
    ::::::::::::::::::::|
DB 283 GGTGCAAGATATGCTCTTCTGTAACGTGCTGATGCGCGCGAAGTGAACCTGTCAGCT 342
    ::::::::::::::::::::|
OY 84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValAlaSerThr 103
    ::::::::::::::::::::|
DB 343 CCAAGTGTCTCTCAAGTGCGCAAAAGAACTAGCTGATTAACAGTGTGTTGTAACACGT 402
    ::::::::::::::::::::|
OY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
    ::::::::::::::::::::|
DB 403 CCATTTACATTTAGAGACGCTGAAGCGTCAACACAGCGCATCTGTTGCGGCAATT 462
    ::::::::::::::::::::|
OY 124 AlaLysTyrValAspThrLeuIleValAlaProAsnGlnAsnLeuLeuAlaAlaAsp 143
    ::::::::::::::::::::|
DB 463 AAAGAAATGTAGTACACTTATGTAATTCACAAAGTATTCGCTTATTAAGATTTGTTGAT 522
    ::::::::::::::::::::|
OY 144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
    ::::::::::::::::::::|
DB 523 AAAAACAAGCCCAATGTTAGAGCATTCGCTGAAGCGGATTAACGATTAACGTCAGAGTGT 582
    ::::::::::::::::::::|
OY 164 LysGlyValThrAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
    ::::::::::::::::::::|
DB 583 CAAGTATTTCTGATTTAATTGCAACACAGGTTAATTAACCTTA 627
    ::::::::::::::::::::|

```

RESULT 7
US-10-282-122A-9479
Sequence 9479, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282.122A
PRIOR FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26

```

PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9479
LENGTH: 1170
TYPE: DNA
ORGANISM: Bacillus anthracis
US-10-282-122A-9479

```

Alignment Scores:

Pred. No.:	8.57e-48	Length:	1170
Score:	505.00	Matches:	100
Percent Similarity:	74.29%	Conservative:	30
Best Local Similarity:	57.14%	Mismatches:	45
Query Match:	57.06%	Indels:	0
DB:	9	Gaps:	0

US-09-770-509-2 (1-178) x US-10-282-122A-9479 (1-1170)

```

OY 4 LeuGlUGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
    ::::::::::::::::::::|
DB 103 GTCAAGAGGTAGACTTTATCGCTGTAATTCAGATGCAACAGCATTAATCATCAAAA 162
    ::::::::::::::::::::|
OY 24 AlaProHIsLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyLysSerLys 43
    ::::::::::::::::::::|
DB 163 GCTGAACAGAAATGCAAAATGCTGCGAATAAAGTAAGAACATTCAGACACTTCCT 222
    ::::::::::::::::::::|
OY 44 ProGluLeuGlyLysArgSerAlaGluGlnGlyValAspIleGlnArgMetLeuGln 63
    ::::::::::::::::::::|
DB 223 CCTGAAGTAGAGAAAGAAAGCTGCGAATAAAGTAAGAACATTCAGACACTTCCT 282
    ::::::::::::::::::::|
OY 64 AspSerAsnMetLeuPheIleThrLysGlyMetGlyGlyThrCysThrGlyAlaAla 83
    ::::::::::::::::::::|
DB 283 GGTGCAAGATATGCTCTTCTGTAACGTGCTGATGCGCGCGAAGTGAACCTGTCAGCT 342
    ::::::::::::::::::::|
OY 84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValAlaSerThr 103
    ::::::::::::::::::::|
DB 343 CCAAGTGTGCTCAAGTGCGCAAAAGAACTAGCTGATTAACAGTGTGTTGTAACACGT 402
    ::::::::::::::::::::|
OY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
    ::::::::::::::::::::|
DB 403 CCATTTACATTTGAGAGCATTCGCTGAACCAACAGCGCATCTGATTAACGCGCATTT 462
    ::::::::::::::::::::|
OY 124 AlaLysTyrValAspThrLeuIleValAlaProAsnGlnAsnLeuLeuAlaAlaAsp 143
    ::::::::::::::::::::|
DB 463 AAAGAAATGTAGTACACTTATGTAATTCACAAAGTATTCGCTTATTAAGATTTGTTGAT 522
    ::::::::::::::::::::|
OY 144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
    ::::::::::::::::::::|
DB 523 AAAAACAAGCCCAATGTTAGAGCATTCGCTGAAGCGGATTAACGATTAACGTCAGAGTGT 582
    ::::::::::::::::::::|
OY 164 LysGlyValThrAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
    ::::::::::::::::::::|
DB 583 CAAGTATTTCTGATTTAATTGCAACACAGGTTAATTAACCTTA 627
    ::::::::::::::::::::|

```

RESULT 8
US-10-282-122A-17147
Sequence 17147, Application US/10282122A
GENERAL INFORMATION:

```

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17147
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Clostridium difficile
US-10-282-122A-17147

Alignment Scores:
Pred. No.: 8,69e-48 Length: 1182
Score: 505.00 Matches: 99
Percent Similarity: 75.71% Conservative: 35
Best Local Similarity: 55.93% Mismatches: 43
Query Match: 57.06% Indels: 0
Gaps: 0

US-09-770-509-2 (1-178) x US-10-282-122A-17147 (1-1182)
QY 2 SerGlnLeuGlnGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
Db 124 GCTCAACCTAAAGGCGTTGAATTTATATCAGTAATTAACATTAACCAAGCTTTATATACCA 183
QY 22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
Db 184 TCAAAAGCTGAAATTAAGTCAAAATGAGAAATAATTAAGTAAAGCACTTGAGACAGCA 243
QY 42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
Db 244 GCAATTCAGAGATGAGGAGAAAGAGCTGCTGAGAGAGAGTAAAGGTGAATTAAGTTA 303
QY 62 LeuGlnAspSerAsnMetLeuPheIleThrGlyMetGlyLysGlyLysThrCysThrGly 81
Db 304 CTTCAGAGAGAGTATGCTTTGTAAACAGCGGATGCGGTGCTACTGCTAACAGCA 363
QY 82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyLysLeuThrValGlyValAla 101

```

```

Db 364 GCTGCTCCAGTGGTTGCAGCGCTTGCCAAAGAAATGGGATCACTTACAGTAGCTAGCTT 423
QY 102 SerThrProPheArgSerGlnGlyProAsnArgThrArgLeuAlaAsnIleGlyValLys 121
Db 424 ACCAAGCTTTTGGATTGTAAGGTAAAGTAAAGTAAAGACGCGTGAAGGTGAATAGCA 483
QY 122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuAlaLeu 141
Db 484 GAACCTAAATCAAAAGTAGATACCTCTTATACGATACCAAAATGATAGACTTTTACAGATA 543
QY 142 AlaAspLysSerThrThrMetLeuGlnAlaPheArgTyrAlaAspAspValLeuGlu 161
Db 544 GTACAAAATAATATACATGTTAGTGCCTTTGCGAGTTCGATGATGATATTAACAACA 603
QY 162 GlValLysGlyValThrAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
Db 604 GGTATACATCAATTTTCAGACTTATACAGAGTTGAAGGTTAATAACTTG 654

RESULT 9
US-10-282-122A-6285
; Sequence 6285, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6285
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-10-282-122A-6285

Alignment Scores:
Pred. No.: 1.29e-46 Length: 1230
Score: 495.00 Matches: 100
Percent Similarity: 73.14% Conservative: 28
Best Local Similarity: 57.14% Mismatches: 47
Query Match: 55.93% Indels: 0

```


DB:	9	Gap:	0
US-09-770-509-2 (1-178) x US-10-282-122a-6285 (1-1230)			
Oy	4	LeuGIuGIyValIGluPheIIeValAlaasnThrAspCyGlnAlaLeuGIyArgSerLeu	23
Db	106	GTAAAGCGCGTGAATTTATCATCACACCAATACAGACGTTCCAAAGCTTTAAACATTCAAAA	165
Oy	24	AlaProHIsLysIleThrLeuGIyLysAspIleThrLysGlyLeuGIyAlaGlySerLys	43
Db	166	GCAGAAACAGATGATTCATTAAGCCCTAAATACACTCGTGTTAGTGCGGGTTCACAA	225
Oy	44	ProGluLeuGIyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln	63
Db	226	CCTGAAGTTGGCCAAAAGCTCCAGAAAGAAAGTCAAGTGAATTCAGAAATCATTAACA	285
Oy	64	AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyThrCysThrGlyAlaAla	83
Db	286	GGCGCGGATATGATTTCTATTCTCTGCTGATGAGGCGGGAACCTGTACAGGTCTGCG	345
Oy	84	ProValValAlaSerValAlaArgGluLeuGIyIleLeuThrValGlyValValSerThr	103
Db	346	CCAGATGCTGCAAAATACGCTAAAGAAATTAAGCGCTTTAAAGTGTGGTGAATCTCGT	405
Oy	104	ProPheArgSerGluGIyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu	123
Db	406	CCATTTAGTTTGAAGGTCCAAAACGCTGCTGCTTTGCTGCTGAAGCAATTCGCTTAATTA	465
Oy	124	AlaLysTyValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp	143
Db	466	AAAGAAACGCTGATACACTATTAATTAATCTCAATTAACCGCTTAATTAAGACGCTTGAC	525
Oy	144	LysSerThrThrMetLeuGlnAlaPheArgTyValAspAspValLeuLeuGluGIyVal	163
Db	526	AAGAAACGCCCAATCTGTGAACCAATTAAGAGAAGCTGATTAATGATTACGTAACGTTGT	585
Oy	164	LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu	178
Db	586	CAAGGATTTTCAGATTTAATCATCTGCACACGAGTTTACGTAAACCTTG	630
RESULT 10			
US-09-134-000C-1253			
GENERAL INFORMATION:			
APPLICANT: Lynn Doucette-Stamm et al			
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO			
FILE REFERENCE: 032796-032			
CURRENT APPLICATION NUMBER: US/09/134, 000C			
CURRENT FILING DATE: 1998-08-13			
PRIOR APPLICATION NUMBER: US 60/055, 778			
PRIOR FILING DATE: 1997-08-15			
NUMBER OF SEQ ID NOS: 6812			
SOFTWARE: PatentIn version 3.1			
SEQ ID NO 1253			
LENGTH: 1308			
TYPE: DNA			
ORGANISM: Enterococcus faecalis			
US-09-134-000C-1253			
Alignment Scores:			
Pred. No.:	1.4e-46	Length:	1308
Score:	495.00	Matches:	100
Percent Similarity:	73.14%	Conservative:	28
Best Local Similarity:	57.14%	Mismatches:	47
Query Match:	55.93%	Indels:	0
DB:	6	Gaps:	0
US-09-770-509-2 (1-178) x US-09-134-000C-1253 (1-1308)			
Oy	4	LeuGIuGIyValIGluPheIIeValAlaasnThrAspCyGlnAlaLeuGIyArgSerLeu	23
Db	181	GTAAAGCGCGTGAATTTATCATCACCAATTAAGACGTTTCAAGCAATTAACATTCAAAA	240

```

QY      24  ALaPrOHIsLySIeThrLeuGIyLysAspRIeThrValGILeuGIyAlaGLysSerIys 43
Db      241  GCAGAAACAGCGATTCATTAGGCCCTTAATACACTCGIGGTITTAGTGCCGGTTACAA 300
QY      44  ProGluLeuGIyLysArgSerIaGLuGIyValAspRIeGLInArgMeLeuGIIn 63
Db      301  CCTGAAGTTGGCCAAAAGAGTCCAGAGAAGAAAGTCAACAAGTATTCAGAAATCATTACAA 360
QY      64  AspSerAasMetLeuPheIIeThrGLyGMetGIyGLyVlThrcysThrGIyAlaIle 83
Db      361  GCGCGCGATGATTCATTTCATCTCGTGTAGGGGGGGAACGCTACAGGTGCTGGC 420
QY      84  ProValIAlaISerValAlaArgSLyuGIyIleLeuThValGILyValIaSerThr 103
Db      421  CCAGTAGTGCAGAAAATGCTTAAGAAATTAGGCCGCTTTAACAGTTGGTGTAGTAACTCG 480
QY      104  ProPheArgSerGIyProAsnArgThrArgLeuAlaAsnAlaGLyValLysGIuLeu 123
Db      481  CCATTAGTATTGTAAGGCTCCAAAACGTGGTGGCTTTGGCTGCAGAGGAATTGCCATTATA 540
QY      124  AlaLysTyrrValAspThrLeuIleValValProAsnGLInAsnLeuAlaLeuAlaAsp 143
Db      541  AAAGAAAACGTGATACACTATTAATTCACAANAACCGCTATTAGAAAGCTGTGAC 600
QY      144  LysSerThrThrMetLeuGIuAlaPheArgTyrrAlaAspAspValLeuLeuGIyVal 163
Db      601  AAGAAACGCCAATGCTTAGAAGCATTAGAGAAGCGATTAAGTATTAGCTAAGGTGT 660
QY      164  LysGLyValThrAspLeuIleValArgProGIyLeuIleAsnLeu 178
Db      661  CAAGGATTCAGATTATCATCTGCACCAGGTGTACGTAAACTTG 705

RESULT 11
US-09-134-000C-1253
; Sequence 1253, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134, 000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1253
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-1253

Alignment Scores:
Pred. No.:          1,4e-46         Length:       1308
Score:              495.00          Matches:       100
Percent Similarity: 73.14%          Conservative: 28
Best Local Similarity: 57.14%        Mismatches:   47
Query Match:        55.93%          Indels:        0
DB:                 6                Gaps:          0

US-09-770-509-2 (1-178) x US-09-134-000C-1253 (1-1308)
QY      4    LeuGIyValIGluPheIIeValAlaAsnThrAspCyGLInAlaLeuGIyArgSerLeu 23
Db      181  GTTAAAGCGCGATTTATTCACAGCCATACAGACGTTCMAAGCTTAATAACATTCAAAA 240
QY      24  ALaPrOHIsLySIeThrLeuGIyLysAspRIeThrValGILeuGIyAlaGLysSerIys 43
Db      241  GCAGAAACAGCGATTCATTAGGCCCTTAATACACTCGIGGTITTAGTGCCGGTTACAA 300
QY      44  ProGluLeuGIyLysArgSerIaGLuGIyValAspRIeGLInArgMeLeuGIIn 63

```

```

Db 301 CCTGAAGTGGCCAAAAGCTGCAGAAAGTGAACAGTATTCAGAAATCATTACAA 360
OY 64 AspSerAsmMetLeupheilethrglylmetglyglythrcysThrGlyAlaIa 83
Db 361 GCGCGGATATGATTTTCTACTGCTGCTGATGGGTGGGACGCTGACGCTGCTG 420
OY 84 ProValValAlaSerValAlaArgGluLeuGlylleuThrValGlyValAlaSerThr 103
Db 421 CCAGTAGTGCAGAAAATCCCTAAAGATTAAGGCTTAAACAGTGGGTGATACCTG 480
OY 104 ProphearserglucylproasnargthrarGluAlaasnAlaGlyValylsleu 123
Db 481 CCATTGATGTTGAGCTCCAAAACGGTGGTTCCTGCTGGAAGGATGGCTTATTA 540
OY 124 AlalystyValAspThrLeuileValAlaProasnGlnAsnLeuAlaLeuAlaasp 143
Db 541 AAGAAAGCTTCATACACTTATATATCTCAATTAACCCCTTATTGAAAGCTTGAC 600
OY 144 LysSerThrThrmleuGluAlaApheargtyrAlaaspValleuLeuGlyVal 163
Db 601 AAGAAAGCCCAATGCTTGAAGCATTTAGAGAACCTGATATGATTAACGCAAGGTGT 660
OY 164 LysGlyValThraspleuileValArgProGlyleuIleasnleu 178
Db 661 CAAGGATTCAGATTAACTCACTGCACGAGTTCAGTAACTTG 705

RESULT 12
US-10-417-884-3641
; Sequence 3641, Application US/10417884
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/417,884
; FILING DATE: 17-Apr-2003
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariunello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3641:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1245 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:

```

```

; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...1245
; SEQUENCE DESCRIPTION: SEQ ID NO: 3641:
US-10-417-884-3641

Alignment Scores:
Pred. No.: 1,71e-46 Length: 1245
Score: 494.00 Matches: 97
Percent Similarity: 73.14% Conservative: 31
Best Local Similarity: 55.43% Mismatches: 47
Query Match: 55.82% Indels: 0
DB: 8 Gaps: 0

US-09-770-509-2 (1-178) x US-10-417-884-3641 (1-1245)

OY 4 LeuGluGlyValAlaGluPheileValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
Db 109 GTTAAAGGTGTTGAATTCATTACAGCCAAACAGAGCTGCAAGCATTAATAAATTCAAA 168
OY 24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
Db 169 GCTGAAACAGTTATCCAAATTAGGCCCCCAAAATATCTCGGTGTTAGTGTCTGTTCTAG 228
OY 44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
Db 229 CCAGAAAGTTGACAAAAGGCGCAGAAAGAACTGAAACATCTTACGCAAGCTTTAGAC 288
OY 64 AspSerAsmMetLeupheilethrglylmetglyglythrcysThrGlyAlaIa 83
Db 289 GGTGACATATATGATTTTCATCATGTCAGAGAAATGGTGGAGAACCGGTACAGTCTGCA 348
OY 84 ProValValAlaSerValAlaArgGluLeuGlylleuThrValGlyValAlaSerThr 103
Db 349 CCAATCGTTCAGATATCGCAAGAAATGGCGGATGACAGTGTGTCGACACGT 408
OY 104 ProphearserglucylproasnargthrarGluAlaasnAlaGlyValylsleu 123
Db 409 CCATTGCTTTGAGAGCCAAACGTCGCTTTGCTGCGAAGGTATCCGCTTTTG 468
OY 124 AlalystyValAspThrLeuileValAlaProasnGlnAsnLeuAlaLeuAlaasp 143
Db 469 AAGAAAGATGTCATATTTATATATTCATTAACCCCTTATTAGACGTTGAGAT 528
OY 144 LysSerThrThrmleuGluAlaApheargtyrAlaaspValleuLeuGlyVal 163
Db 529 AAAAAGACACCGATGCTTGAAGCTTCCGTGAAGCAGATTAAGTTTACGACAAAGGTGA 588
OY 164 LysGlyValThraspleuileValArgProGlyleuIleasnleu 178
Db 589 CAAGTATCTCTGATTTGATCAGACGCGCAGGTTACGTTAACTT 633

RESULT 13
US-10-282-122A-21408
; Sequence 21408, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20

```

```

; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 21408
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Enterococcus faecium
; US-10-282-122A-21408

Alignment Scores:
Pred. No.: 1,71e-46 Length: 1245
Score: 494.00 Matches: 97
Percent Similarity: 73.14% Conservative: 31
Best Local Similarity: 55.43% Mismatches: 47
Query Match: 55.82% Indels: 0
DB: 9 Gaps: 0

US-09-770-509-2 (1-178) x US-10-282-122A-21408 (1-1245)
QY 4 LeuGlUGlYAlGluPhelIeValAlAsnThrAspCysGlnAlaLeuGlYArgSerLeu 23
DB 106 GTTAAAGCTGTGATTCATTACACGCCAACACAGCGTCAAGCATTAATAAATTCAAA 165
QY 24 AlaProHslYsIleThrLeuGlYAspIleThrYsGlyLeuGlYAlaGlySerLys 43
DB 166 GCTGAACAGTATTCATTCATAGCCCAATATACCTCGTGTAGGTGCTGCTTCACG 225
QY 44 ProGluLeuGlYAspSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
DB 226 CCAAGAACTGGACAAAAGCGGACGAAAGAAAGTGAACATCCCTTACGCGAAGCTTGAAC 285
QY 64 AspSerAsnMetLeuPhelIeThrGlyMetGlyGlyThrCysThrGlyAlaAla 83
DB 286 GGTGACATATGATTTTCATCAGCAGCAATGGGTGGAGAACCGGTACAGTCTGCTCA 345
QY 84 ProValValAlaSerValAlaArgGluLeuGlYIleLeuThrValGlyValAlaSerThr 103
DB 346 CCAATCGTTCAGCATTCGCAAGAAATTTGGCGCATTCAGTGTGGTGTGCGTACGCT 405
QY 104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
DB 406 CCAATTCCTTTGAGGACGCAAAACGTCGCTTTGTCGCAAGAAAGTATCGCTCGTTG 465
QY 124 AlaLysrYrValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaAsp 143
DB 466 AAGAAATATGCGTACATTTATATTCATAATTAACCGCTTATATAGAACTGTAGAT 525
QY 144 LysSerThrThrMetLeuGluAlaPheArgTyTAlaAspAspValLeuLeuGluGlyVal 163
DB 526 AAAAAGACACCGAGCTTGAAGCTTTCGCTGAAAGCAGATTAATGTTTACGACAAAGCTGA 585
QY 164 LysGlyValThrAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
DB 164 LysGlyValThrAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
```

```

DB 586 CAAGTATCTCTGATTTGATCAGCGCCAGGTACCTTACCTT 630
RESULT 14
US-10-282-122A-25784
; Sequence 25784, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 25784
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Mycobacterium avium
; US-10-282-122A-25784

Alignment Scores:
Pred. No.: 2.64e-46 Length: 1161
Score: 492.00 Matches: 96
Percent Similarity: 74.86% Conservative: 35
Best Local Similarity: 54.86% Mismatches: 44
Query Match: 55.59% Indels: 0
DB: 9 Gaps: 0

US-09-770-509-2 (1-178) x US-10-282-122A-25784 (1-1161)
QY 4 LeuGlUGlYAlGluPhelIeValAlAsnThrAspCysGlnAlaLeuGlYArgSerLeu 23
DB 97 CTCGAAGGCGGTGAGTTCATTCGCGATCAACACCGCGCAAGCGCTGTGATGAGCAT 156
QY 24 AlaProHslYsIleThrLeuGlYAspIleThrYsGlyLeuGlYAlaGlySerLys 43
DB 157 GCCGAGCTCAAACTCAGCTCGGCGCGGACTCCACCGCTGAGCTGGCGCGCGGAC 216
QY 44 ProGluLeuGlYAspSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
DB 217 CCGGAGGTGCGCGGCGCGGCGCGGAGGACCCCAAGACGATCGAGAGCTGCTGCGC 276
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 2, 2003, 08:47:53 ; Search time 1762 seconds

(without alignments)
1636.093 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASOLEGVEFIYANTDCALG.....LLEGVGYDLIVRGLINL 178

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q/cgcn2.1/USFTO.spool/US09770509/runat_27052003_083649_18832/app.query.fasta_1.327
-DB=EST -OFMT=fastlap -SUFFIX=est -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCCALIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcr -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09770509.cgcn2.1.1906 @runat_27052003_083649_18832 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MAMP -LARGQUERRY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlun:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hnc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hnc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	449	50.7	570	9 A1775094	A1775094 EST756194
2	438	49.5	1562	11 A1709745	A1709745 Zea mays
3	435	49.2	668	14 BQ578606	BQ578606 WHE0307_E
4	433	48.9	540	9 A1469754	A1469754 A1469754
5	427	48.0	593	10 BE498211	BE498211 WHE0954_C
6	425	48.2	1537	11 A1106558	A1106558 Zea mays
7	412	46.6	627	10 A1775962	A1775962 EST735027
8	404	44.6	664	12 BG855721	BG855721 1024043F0
9	395	44.6	670	10 AM094550	AM094550 EST287730
10	379	42.8	621	13 B1289183	B1289183 B1289183
11	368	41.6	615	13 B1361253	B1361253 B1361253
12	363	41.0	661	13 B1415730	B1415730 B1415730
13	360	40.7	663	14 BQ279894	BQ279894 1091033D0
14	358	40.5	765	14 BQ012384	BQ012384 OGJ1M13.Y
15	352	39.8	765	14 BQ869798	BQ869798 QGD7D11.Y
16	351	39.7	599	14 BQ865867	BQ865867 QGC6B17.Y
17	351	39.7	731	14 BQ855906	BQ855906 GGR28A23
18	350	39.5	654	10 A1934672	A1934672 EST353564
19	347	39.2	598	13 B1787678	B1787678 sal148d09
20	346	39.1	426	13 B1942496	B1942496 sl64d07.Y
21	337	38.1	409	17 A2302940	A2302940 GSSBRu188
22	336	38.0	541	13 B1463421	B1463421 B1463421
23	332	37.5	608	13 B1325414	B1325414 B1325414
24	331	37.4	605	13 B1329349	B1329349 B1329349
25	329	37.2	544	12 BE941644	BE941644 EST421223
26	325.5	36.8	752	14 BQ011950	BQ011950 OGJ17324
27	319	36.0	492	10 A1559837	A1559837 A1559837
28	316	35.7	538	10 BE444075	BE444075 WHE1127_G
29	311	35.1	404	10 A1624700	A1624700 A1624700
30	308	34.8	458	13 B1930248	B1930248 EST550137
31	303.5	34.3	641	17 B1397494	B1397494 AG-ND-119
32	300	33.9	532	14 BQ279544	BQ279544 1091036G1
33	298.5	33.7	550	17 B1397266	B1397266 AG-ND-119
34	294	33.2	467	17 A2049166	A2049166 GSSBRu056
35	294	33.2	655	14 BQ402273	BQ402273 GA_EB004
36	292	33.0	814	10 BE643351	BE643351 CT12_8_J1
37	290	32.8	448	17 B1751621	B1751621 SALK_0503
38	289	32.7	609	10 BE343445	BE343445 EST408607
39	288	32.5	455	17 B1751623	B1751623 SALK_0503
40	288	32.5	495	10 BE025134	BE025134 B94008C08
41	287	32.4	544	13 B1324543	B1324543 B1324543
42	273	30.8	468	12 BG510632	BG510632 gsc71403
43	271	30.6	431	17 A0863118	A0863118 nbebd0020E
44	266.5	30.1	723	12 BG681299	BG681299 1024076G1
45	263	29.7	583	10 AV916706	AV916706 AV916706

ALIGNMENTS

RESULT 1
LOCUS A1775094
DEFINITION EST756194 tomato resistant, Cornell Lycopersicon esculentum cDNA
ACCESSION A1775094
VERSION A1775094.1 GI:5273135
KEYWORDS
SOURCE
ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 570)


```

Db      535  ACCTATTCATTGCTTCAGGCGCCGTAAAGCCCTCTGACAGCATGTGACACTAGAG 594
Oy      122  GULEUAlaLysrYrValaPThrLeuIleValaProaAnglnAsnLeuAlaLeu 141
Db      595  AACCTGAAAGAGTGTACACACTTATGTGTATCCAAATGATAACTATTATGATGT 654
Oy      142  AlaAspLysSerThrTrpMetLeuGluAlaPheArgTyAlaAspAspValLeuGlu 161
Db      655  GCGGATGAAGAACATGCCCTTGCAAGATCATTTCTCTCCAGATGATGCTCTGTCAG 714
Oy      162  GlyValIleGlyValIleThrAspLeuIleValaArgProGlyLeuIleAsnLeu 178
Db      715  GGTGTTCAGAAATATCATGACATCATCATCAATACCGGACTGTGTCAATGTT 765

RESULT 3
BOS78606 668 bp mRNA linear EST 19-JUN-2002
LOCUS WHE0307_E03_J0525 wheat unstressed seedling shoot cDNA library
DEFINITION Triticum aestivum cDNA clone WHE0307_E03_J05, mRNA sequence.
ACCESSION BOS78606
VERSION BOS78606.1 GI:21481923
KEYWORDS EST
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
1 (bases 1 to 668)
Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han
, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Raush, C.J.,
Seaton, C.L. and Tong, J.C.
The structure and function of the expressed portion of the wheat
genomes - Etisolated shoot cDNA library
Unpublished (2000)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: SK primer.
Location/Qualifiers
1..668
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE0307_E03_J05"
/tissue_type="Wheat unstressed seedling shoot cDNA library"
/dev_stage="Five day old seedling"
/lab_host="E. coli SOLA"
/Note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized
, germinated and grown aseptically in the dark at room
temperature on filter paper with water, nystatin and
cefotaxime in covered crystallization dishes. Shoots were
harvested. The tissue, total RNA, and poly(A) RNA were
prepared, a cDNA library was made, and the cDNA clones
were in vivo excised to give Bluescript phagemids in the
T3 Close lab (Choi, Close, Fenton) at the University of
California, Riverside. Plasmid DNA preparations and DNA
sequencing were performed in the OD Anderson lab (all
other authors)."
BASE COUNT 159 a 137 c 196 g 176 t
ORIGIN
Alignment Scores: 9,66e-41 Length: 668
Pred. No.: 435.00 Matches: 81
Score:

```

```

Percent Similarity: 68.33% Conservative: 42
Best Local Similarity: 45.00% Mismatches: 55
Query Match: 49.15% Indels: 2
DB: 14 Gaps: 1

US-09-770-509-2 (1-178) x BOS78606 (1-668)

Oy      1  AlaSerGlnLeuGluGlyValGluPheIleValaAlaAsnThrAspCysGlnAlaLeuGly 20
Db      7  GCACGAGCATGATGATGTCGAGATTTGGATGTCACACACCGATGTCACAGGCATAAAG 66
Oy      21  ArgSer-----LeuAlaProHisLysIleThrLeuGlyLysAspIleThrGlyGlyLeu 38
Db      67  ATGTCCCGCGGTGATTCGCCAGACAGGCTCAGATGGCAGAGCTCCTCGGGGTTTG 126
Oy      39  GlyAlaGlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValaAspIle 58
Db      127  GGTGCGGTGGGAGAACCTGATATTGGGATGAATGCCCAAGAGAGCTGTGATCCATA 186
Oy      59  GlnArgMetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyThr 78
Db      187  GAGGAAGCTCTTCATGCTGCTGACATGTTTGTCTACGGCTGGAATGGGTGGGGAAC 246
Oy      79  CysThrGlyAlaAlaAlaProValAlaAlaSerValAlaArgGluLeuGlyIleLeuThrVal 98
Db      247  GGAAGCTGAGAGTGCCTCCCTGTAATGCTGGAATGCCAAGATCCAGATGATTAACAGTG 306
Oy      99  GlyValAlaSerThrProPheArgSerGlyGlyProAsnArgThrArgLeuAlaAsnAla 118
Db      307  GGTATGTTCACAACGCCCTTTCATTTGAGGGAGAGGAGCGGCAAGTTCAGGCTCAGAA 366
Oy      119  GlyValIleGlyLeuLeuAlaLysrYrValaPThrLeuIleValaProaAnglnAsnLeu 138
Db      367  GGAATATCAGCTTGGAAGAAATAGTGTGACACTCTCTTCTCATCCCAATATGACACAGTG 426
Oy      139  LeuAlaLeuAlaAspLysSerThrTrpMetLeuGluAlaPheArgTyAlaAspAspVal 158
Db      427  TTGTCTGCTGTTCTTCCAATATCTCCTGTCACGAGAGCATTAACCTGGGTGATGATATT 486
Oy      159  LeuLeuGlnGlyValIleGlyValIleThrAspLeuIleValaArgProGlyLeuIleAsnLeu 178
Db      487  CTTTGGCAAGAAATTCGTGTATCTGTGATATCATTAACGTTCTGTGGTGTAAATGTT 546

RESULT 4
AJ469754 540 bp mRNA linear EST 24-MAY-2002
LOCUS AJ469754
DEFINITION Hordeum vulgare cDNA clone S000080253A03F1, mRNA
sequence.
ACCESSION AJ469754
VERSION AJ469754.1 GI:21185710
KEYWORDS EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
1 (bases 1 to 540)
; Trilicaceae; Hordeum.
REFERENCE
1 Saren, A.-M., Tanskanen, J., Paulin, L. and Schulman, A.H.
Unpublished (2002)
TITLE Barley EST's
JOURNAL
COMMENT Contact: Schulman AH
Institute of Biotechnology
University of Helsinki
P.O.Box 56 (Valkankari 6A), University of Helsinki FIN-00014,
Finland.
FEATURES
SOURCE
1..540
Location/Qualifiers
/organism="Hordeum vulgare"
/db_xref="taxon:4513"
/clone="S000080253A03F1"
/tissue_type="S00008"
/Note="Callus K19"

```


Db	510	ATATCAGACATTATC	524
RESULT 6	AY110658	1537 bp	mRNA linear HTC 26-MAY-2001
LOCUS	AY110658		
DEFINITION	Zea mays CL806_1 mRNA sequence.		
ACCESSION	AY110658		
VERSION	AY110658.1	GI:21215248	
KEYWORDS	HTC.		
SOURCE	Zea mays.		
ORGANISM	Zea mays.		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 1537) Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.		
TITLE	Maize Mapping Project/Dupont Consensus Sequences for Design of Overgo Probes		
JOURNAL	Unpublished (2002)		
REFERENCE	2 (bases 1 to 1537) Coe,E.C.		
AUTHORS	Direct Submission		
TITLE	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA		
JOURNAL	Location/Qualifiers		
FEATURES	1..1537 /organism="Zea mays" /db_xref="MaizeDB:632908" /db_xref="taxon:457" /clone="CL806.1" /clone_lib="Maize Mapping Project/Dupont Consensus Library" /note="this sequence is part of a project of EST assemblies resulting from the application of public configs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"		
BASE COUNT	352 a 311 c 404 g 374 t 96 others		
ORIGIN			
Alignment Scores:			
Pred. NO.:	4.33e-39	Length:	1537
Score:	425.00	Matches:	80
Percent Similarity:	66.48%	Conservative:	39
Best Local Similarity:	44.69%	Mismatches:	58
Query Match:	48.02%	Indels:	2
DB:	11	Gaps:	1
US-09-770-509-2 (1-178) x	AY110658 (1-1537)		
QY	2	SerGlnLeuGluGlyValGlnPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg	21
Db	185	AGCTTCACATGAACGGCGTGAAGTTTGGATCGCGAACACTGATGTCGACGCCATTAAGATG	244
QY	22	Ser-----LeuAlaProHisLeuIleThrLeuGlyLysAspIleThrLysGlyLeuGly	39
Db	245	TCTCTGTCGCTNNNNNNACAAATAGACTCGAATTTGACAGAGAACTGATGATCGAGCGCTGGGC	304
QY	40	AlaLeuSerLysProGlnLeuGlyLysArgSerAlaGlnGlnGlnLysValAspIleGln	59
Db	305	GCTNNNNNMAACCCCTGATATTGGGATGTAATGACGCAAAAGGAGAGACGACCATTCAG	364
QY	60	ArgMetLeuGlnAlaSerSerAspMetLeuPheIleThrGlyLysMetGlyValGlyLysThrCys	79
Db	365	GAACCTCNNNNNGGCTGCATATGNNNNNNNGGACGCGCTGGAAATGGGTGGAGGAACGGA	424
QY	80	ThrGlyAlaAlaProValValAlaSerValAlaArgGlnLeuGlyIleLeuThrValGly	99
Db	425	ACTGGAGAGCTGCTCGTATGTCGATGATGAGCAAGATGCATGCGGTATACTAACCTGTGC	484
QY	100	ValValSerThrProPheArgSerGlnLysProAsnArgThrArgLeuAlaAsnAlaGly	119

Db	485	ATATCTACACACGCGCTTTCTCGTTGAGGGGGAAGACGGCGAGTTCAAGCTTCACGAGGCA	544
Oy	120	VallysgluLeuAlaLysTYrValAspThrLeuIleValValProAsnGlnAsnLeuLeu	139
Db	545	ATACGACGATGTGAAATATAGTGTGGACACCGCTATTCGTCAATCCCAATGATTAAGTTGCTG	604
Oy	140	AlaLeuAlaAspLysSerThrThrMetLeuGlnAlaPheArgTYrAlaAspAspValLeu	159
Db	605	TCTCTGCTTTCTCCAAATACACCGTGAAGTCAATGATTAATTCGTGCTGATGATTAATCTT	664
Oy	160	LeuGluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu	178
Db	665	CGTCAAGCGATCTGTGGCATATCTGATTAATTAATTCAGCTTCTCGGTGGTTAAATGTT	721
RESULT 7			
LOCUS	AW75962	627 bp	mRNA linear EST 07-SEP-2000
DEFINITION	EST330327	DSIL	Medicago truncatula CDNA clone pDSIL-3J17, mRNA
ACCESSION	AW75962		
VERSION	AW75962.1	GI:7765775	
KEYWORDS	EST.		
SOURCE	barrel medic.		
ORGANISM	Medicago truncatula		
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Medicago.		
AUTHORS	1 (bases 1 to 627) Fedorova,M., Pearson,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Peng,H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and Fraser,C.M.		
TITLE	ESTs from leaves of Medicago truncatula after inoculation with Colletotrichum trifolii		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Deborah A. Samac Department of Plant Pathology University of Minnesota 495 Borling Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA Tel: 612 625 1243 Fax: 651 649 5058 Email: debby@puccln1.crl.umn.edu Minnesota sequence name: M2S9295e TIGR sequence name: MTFAK57TK More information is available at: . http://chryste.tamu.edu/medicago Seq primer: SKmod (CTA gaa cta gta gat CC). Location/Qualifiers 1..627 /organism="Medicago truncatula" /cultivar="genotype A17" /db_xref="taxon:3880" /clone="pDSIL-3J17" /clone_lib="DSIL" /tissue_type="leaves infected with Colletotrichum trifolii" /dev_stage="cotyledons and primary leaves harvested 5 and 8 days after inoculation with Colletotrichum trifolii" /lab_host="E. coli strain X10LR" /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; cDNA was prepared from polyA+ enriched RNA from cotyledons and primary leaves harvested 5 and 8 days after inoculation with Colletotrichum trifolii. The cDNA was directionally ligated into the Uni-ZAP XR vector from Stratagene and packaged using GigaPack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in X10LR cells. Note: EST may be of fungal origin."		
BASE COUNT	166 a	112 c	161 g 187 t 1 others
ORIGIN			

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	source
LYCOPERSICON ESCULENTUM	1 (bases 1 to 670)	D'Ascenzo, M., He, X., Lyman, J., Holt, I. E., Liang, F., Upton, J., Rinning, C. M., Craven, M. B., Fujii, C. Y., Bowman, C. L., Nierman, W., Fraser, C. M., Venter, J. C., Martin, G. B., Tanksley, S. D. and Giovannoni, J.	Generation of ESTs from tomato leaf tissue	Unpublished (1999)	Contact: CUGI Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: http://www.genome.clemson.edu/orders/index.html 5 prime sequence.	Location/Qualifiers	1. 670 /organism="Lycopersicon esculentum" /cultivar="Rio Grande PGR" /db_xref="taxon:4081" /clone="cLET29K5" /clone_lib="tomato mixed elicitor, BRT" /tissue_type="leaf" /dev_stage="4-6 week old plants" /lab_host="XLI-blue MRF" /note="Vector: plasmidscript SK(-); Site_1: EcoRI; Site_2: XhoI; cLET inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisonicotinic acid, BTH, jasmonic acid, ethylene, fenchone, Eix, okadaic acid, or systemin prior to tissue harvest. EcoRI site was destroyed during cloning."
BASE COUNT	187 a	114 c	181 g	188 t			
ORIGIN							
Alignment Scores:							
Pred. No.:	4.67e-36	Length:	670				
Score:	395.00	Matches:	71				
Percent Similarity:	71.24%	Conservative:	38				
Best local Similarity:	46.41%	Mismatches:	44				
Query Match:	44.63%	Indels:	0				
DB:	10	Gaps:	0				
US-09-770-509-2 (1-178) x AW094550 (1-670)							
OY	26	HisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGlu	45				
OY	5	AATGATTTGCCAATATAGCCCATTACTCTCGAGAGACAGTGTCCAGTGGTAATCCAGAT	64				
OY	46	LeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSer	65				
OY	65	ATAGCGATGTAATGCTGCCCAAGAAAGCAGAGAGAGCTATCGAAGAGCGTTCGCGGTCA	124				
OY	66	AsnMetLeuPheIleThrGlnGlyMetGlyGlyGlyThrGlyThrGlnGlyAlaAlaProVal	85				
OY	125	GATATGGTTTGTGACTGCTGGATATGGCGGAGAGAACAGGACTGGGGGCTCTATTA	184				
OY	86	ValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValAlaSerThrProPhe	105				
OY	185	ATTTCAGAGAAATGTCGAATCAATGAGATCTTAATGTTGATATGTCACAAACCCCTTT	244				
OY	106	ArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLys	125				
OY	245	TCCTTTAGGAGCAGACAGAGAGCAGTTCAACCCAGAGAAAGAAATGTCAGCTTATAGAGA	304				
OY	126	TyrValAspThrLeuIleValAlaProAsnGlnAsnLeuAlaLeuAlaAspLysSer	145				
OY	305	AATGTTATACACATATGTCATTCCTTAATGACAAAGTACTGACTGCTGTTCTCTATCA	364				
OY	146	ThrThrMetLeuGlnAlaPheArgTyrTyrAlaAspValLeuLeuGluGlyValLysGly	165				

Db	365	ACCAGCACTACTGAGACCTTTTAACCTGGCTGATCATATATCTTCGGCAAGGGGCTTCGTGT	424
Qy	166	Valnhraspleuilevalargproglyleuileasnleu	178
Db	425	ATTCTGATATTAATTACGATCTCTGAGCTGTAATGT	463
RESULT_10			
LOCUS	BJ289183		
DEFINITION	BJ289183 Y. Ogihara unpublished cDNA library, Wh_SL Triticum aestivum cDNA clone whsl3n10 5', mRNA sequence.		
ACCESSION	BJ289183		
VERSION	BJ289183.1		
KEYWORDS	EST.		
SOURCE	bread wheat.		
ORGANISM	Triticum aestivum		
REFERENCE	Ekarkyoti; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae		
AUTHORS	/ Triticaceae; Triticum.		
JOURNAL	1 (bases 1 to 621)		
COMMENT	Ogihara,Y. and Murai,K. Expressed genes in Triticum aestivum Unpublished (2002) Contact: Tadasu Shin-1 Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshin@genes.nig.ac.jp. Location/Qualifiers		
FEATURES			
source	1..621		
	/organism="Triticum aestivum"		
	/cultivar="Chinese Spring"		
	/db_xref="taxon:4565"		
	/clone="whsl3n10"		
	/clone_id="Y. Ogihara unpublished cDNA library, Wh_SL"		
	/tissue_type="seed DPA30"		
	/dev_stage="Feekes' scale 11.3"		
	/note="Vector: Lambda Uni-ZAP XR, excised phagemid: site.1: EcoRI; site.2: XhoI; plants were grown under hydroponic conditions at UC Davis, salt stressed for 12 hours, and for 7 days, then dissected and frozen (Akhunov in J. Dvorak Lab). Total RNA was prepared from sheath tissue, equal quantiles of RNA were pooled from the two samples, polyA was purified from the pooled RNA, a cDNA library was made, and the cDNA clones were in vivo excised to give phagescript phagemids in the T1 Clonase lab at the University of California, Riverside (Akhunov, Chin , Choi, Clouse, Fenton, Kisanian, Otto, Simons, Zhangy). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."		
BASE COUNT	150 a 120 c 187 g 163 t		
ORIGIN			
Alignment Scores:			
Pred. NO.:	3,15e-34	Length:	621
Score:	379.00	Matches:	69
Percent Similarity:	69.44%	Conservative:	31
Best Local Similarity:	47.92%	Mismatches:	44
Query Match:	42.82%	Indels:	0
Db:	13	Gaps:	0
US-09-770-509-2 (1-178) x BJ289183 (1-621)			
Qy	35	Thrltyscltyeuglyalaglyserlyproglubendglylarsargseralagluginln	54
Db	20	ACACGGGGCTCTGGTGGCGGAGGGAACCCGATTTGGATGATGCTGCCAAGACGAC	79
Qy	55	LyseValaAspIleaglInarqmetleuGlnaAspSerAsnMetleuPheIletrnglyMet	74
Db	80	TGTGAGTCCATTAGAGGAACACTTCAGTCAGTCGTCACATGCTTTTGTACAGGACGAGATG	139

QY	75	GLYILGLYTHR ¹ PCYST ² THGL ³ VAL ⁴ ALA ⁵ PRO ⁶ VAL ⁷ ALA ⁸ SER ⁹ VAL ¹⁰ ALA ¹¹ SER ¹² VAL ¹³ ALA ¹⁴ ARG ¹⁵ LEU ¹⁶ GLU ¹⁷ GLU ¹⁸	94
Db	140	GGTGGGGAACTGGAACTGGAGAGTGGCCCCCTGTAATTGCTGGAAATGGCCAGATGCATGGGT	199
QY	95	ILELEUTHRA ¹ LGLY ² VAL ³ VAL ⁴ SER ⁵ THR ⁶ PRO ⁷ PHEAR ⁸ GER ⁹ GLUGLY ¹⁰ PRO ¹¹ ASN ¹² ARG ¹³ THR ¹⁴ ARG ¹⁵	114
Db	200	ATATCTGCACAGGGGATATTGTCACCAAGCCCTTTTCATTATTTGGGGGAGGAGAGCGGGCAGTT	259
QY	115	LEU ¹ ALASNA ² LGLY ³ VAL ⁴ GLY ⁵ GLU ⁶ VAL ⁷ GLY ⁸ TYR ⁹ VAL ¹⁰ ASP ¹¹ THR ¹² LEU ¹³ ILE ¹⁴ VAL ¹⁵ PRO ¹⁶	134
Db	260	CAGGCTCAACAGGAAGAAATATTCAGCCTTGAGAAATAGTGGAGACACTCTCATTTGTCATCCCA	319
QY	135	ASNGLINS ¹ LEU ² LEU ³ ALA ⁴ LEU ⁵ ALA ⁶ ASP ⁷ YSER ⁸ THR ⁹ THR ¹⁰ MEL ¹¹ LEU ¹² GLU ¹³ ALA ¹⁴ PHEAR ¹⁵ GLY ¹⁶	154
Db	320	AATACACAAGCTGTGTGCTGCTGTTCTCCAAACAACACTCTGTACGGAAGCATTTCAACTTG	379
QY	155	ALASPASPA ¹ VAL ² LEU ³ LEU ⁴ GLUGLY ⁵ VAL ⁶ LYS ⁷ GLY ⁸ VAL ⁹ THR ¹⁰ ASP ¹¹ LEU ¹² ILE ¹³ VAL ¹⁴ ARG ¹⁵ PRO ¹⁶ GLY ¹⁷	174
Db	380	GCTGATGATATTTCTTTGGCAGGAATTCGGCGTATCTCTGATATCATATTACGGTTCCTGGG	439
QY	175	LEU ¹ ILEASN ² LEU ³ 178	
Db	440	CTGGTTAATGTT 451	

RESULT 11	
BJ361253	
LOCUS	
DEFINITION	BJ361253 Dictyostelium discoideum cDNA library, CP Dictyostelium
ACCESSION	dictoideum cDNA clone ddc16n12.5', mRNA sequence.
VERSION	BJ361253
KEYWORDS	BJ361253.1 GI:19260848
SOURCE	EST.
ORGANISM	Dictyostelium discoideum.
REFERENCE	Dictyostelium discoideum.
AUTHORS	Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
TITLE	1 (bases 1 to 615)
	Utsunishara,H., Tanaka,Y., Kohara,Y. and Shin-I,T.
	Full length cDNA of Dictyostelium discoideum at the culmination
	stage

JOURNAL Unpublished (2002)
COMMENT Contact: Tadashi Shln-1
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.
FEATURES
 Location/Qualifiers
 1. 615

BASE COUNT	ORIGIN
242 a	1 others
77 c	171 t
124 g	

Alignment Scores:	
Pred. No.:	6, 03e-33
Score:	366.00
Percent Similarity:	72.26%
Best Local Similarity:	51.09%
Query Match:	41.58%
DB:	13
Length:	615
Matches:	70
Conservative:	29
Mismatches:	38
Indels:	0
Gaps:	0

US-09-770-509-2 (1-178) x BJ361253 (1-615)

3 glnleuglgluylalphelelvalalasnthrpsycysglnalaleuylarser 22
 196 GAATTATATGGAATTGATTTTGTGACTTGCCCATACTGATGCACAGCAATTTGGCATATATCA 255

Oy 23 LeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySer 42
 Db 256 TGTACTAGAAAGATGTCACAAATTAGGAAGAACATTACCAAGAGGATTAGAGCGAGGCA 315
 Oy 43 LysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeu 62
 Db 316 GTACCAGAAAGTTGGAAAGAAACCACTGGAAGATCATTTGAAAGCAATTATGATCAAAATT 375
 Oy 63 GlnAspSerAsnMetLeuPheIleThrGlyMetGlyGlyLysThrGlyAla 82
 Db 376 GGTGATACCAAAATGTTATTTGTACACAGCTGATAGGGTGGTGTACAGTACAGSTGGA 435
 Oy 83 AlaProValAlaIleSerValAlaArgGluLeuGlyIleLeuThrValAlaIleValSer 102
 Db 436 GCAGCAGTATTATTCATCCACGACCAAGCAAGCAAGTATTTTAACTGGTGTGTTGTAAC 495
 Oy 103 ThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnIleGlyValLysGlu 122
 Db 496 AAACCATTTTCATTTCGAAAGGTAAACATATGAAATGGCAAGAACCAAGGTTTGATAGAG 555
 Oy 123 LeuAlaLysTyrValAspThrLeuIleValAlaProAsnGlnAsnLeuLeu 139
 Db 556 TTGGGGAATACAGTCGATAGTTTAACTCGTTATTCCAATGAGAAATTAATG 606

RESULT	12
BJ415730	
LOCUS	661 bp mRNA linear EST 10-MAR-2002
DEFINITION	BJ415730 Dictyostelium discoideum cDNA library; VR Dictyostelium discoideum cDNA clone dvv3d18 5' , mRNA sequence.
ACCESSION	BJ415730
VERSION	Dictyostelium discoideum .
KEYWORDS	EST.
SOURCE	Dictyostelium discoideum . Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium. bases 1 to 661)
ORGANISM	Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T. Full length CDNA of Dictyostelium discoideum at the vegetative stage
REFERENCE	
AUTHORS	
TITLE	

JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shii-i
Center for Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES Location/Qualifiers
source 1. .661

```

/organism="Dictyostellium discoideum"
/strain="Ax4"
/db_xref="taxon:44689"
/clone="ddv23d18"
/clone_1ib="Dictyostellium discoideum cdna library, VF
/sex="mat A"
/dev_stage="Growth phase"
BASE COUNT      257 a      87 c      134 g      182 t      1 others
ORIGIN

```

Alignment Scores:	
Pred. No.:	2,55e-32
Score:	363.00
Percent Similarity:	71.53%
Best Local Similarity:	50.36%
Query Match:	41.02%
BB:	13
Gaps:	0
length:	66
Matches:	69
Conservative:	29
Mismatches:	39
Indels:	3
Gaps:	0

US-09-770-509-2 (1-178) x BJ415730 (1-661)

Qy 3 G L T L e u g L u g L y a l g l u p h e i l e v a l A l a s n t h A s p c y s g l n a l e u l y a r g s e r 22
 ||||| |||||
 Db 241 G A A T T A T A T G G A A T T G A T T T G T A C T T G C C A T A C T A C G A T C A C A A G C A A T T G G C A A T T C A 3000

QY 23 LeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuAlaGlySer 42
 DB 301 TGTAGTAGAAGATGTGACATATAGAAAGACATTAACAGAGGATAGACAGAGGACA 360
 QY 43 LysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeu 62
 DB 361 GATACGAAAGTTGAAAGAAAGCACTGACAAATCAATTAAGAAATTAATCAAAAT 420
 QY 63 GlnAspSerAsnMetLeuPheIleThrGlyLysGlyGlyGlyLysThrCysThrGlyAla 82
 DB 421 GGTGATACACAAATGTTATTTGTACACAGCTGATGGTGTGTACAGAGTACAGTGGCA 480
 QY 83 AlaProValAlaAlaSerValAlaArgGluLeuGlyLysLeuThrValGlyValAlaSer 102
 DB 481 GCACGAGTATTTGATCATCAGCAGCAAAAGCCAAAGGATTTTAACTGTTGATTTGAACC 540
 QY 103 ThrProPheArgSerGlyGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGlu 122
 DB 541 AAACCATTTTCATTTTCGAAAGTAAACATGAAATGGAATGGCAACAAAGTTGATATAGAG 600
 QY 123 LeuAlaLysTyrValAlaSerThrLeuIleValAlaProAsnGlnAsnLeuLeu 139
 DB 601 TTGGAAGAAATCAGNCAGTACTTATCTTATTCCTCAATGAGAAATTAATG 651
 RESULT 13
 LOCUS BQ279894 663 bp mRNA linear EST 08-MAY-2002
 DEFINITION 1091033005.y1.1091 - Immature ear with common ESTs screened by
 Schmidt Lab Zea mays cDNA, mRNA sequence.
 ACCESSION BQ279894
 VERSION BQ279894.1 GI:20508091
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 663)
 Walbot, V.
 Maize ESTs from various cDNA libraries sequenced at Stanford
 University
 Unpublished (1999)
 CONTACT: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2327
 Fax: 650 723 8221
 Email: walbot@stanford.edu
 Plate: 1091033 row: D column: 05.
 FEATURES
 source
 1..663
 /organism="Zea mays"
 /cultivar="OH43"
 /db_xref="taxon:4577"
 /clone_lib="1091 - Immature ear with common ESTs screened
 by Schmidt Lab"
 /tissue_type="Inflorescence meristem - floral organ
 primordia"
 /dev_stage="0.5 cm to 2 cm"
 /lab_host="Stratagene XL0LR"
 /note="Organ: Immature ear; Vector: pAD-GAL4; Site: 1;
 EcoRI; Site 2: XhoI; RNA from library 606 was filtered for
 common ESTs found in 605."
 BASE COUNT 147 a 172 c 196 g 148 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.75e-32 Length: 663
 Score: 360.00 Matches: 72
 Percent Similarity: 70.83% Conservative: 30
 Best Local Similarity: 50.00% Mismatches: 42
 Query Match: 40.68% Indels: 0

DB: 14 Gaps: 0
 US-09-770-509-2 (1-178) x BQ279894 (1-663)
 QY 2 SerGlnLeuGluGlyValGlnPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
 DB 217 AGCGGCTCCAGGGGATGCAATTTATGCTATTAACACCATTCGCCAAGCCCTTATTAAT 276
 QY 22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuAlaGly 41
 DB 277 TCACAGGCGCAATATCTCTGCAAAATGGAGACAGTGTACCGCGCTTAGGTAGTGTGT 336
 QY 42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
 DB 337 GGAATCCGAATTTGGAGAGAGAGCGCTGCTGAGGAATCAAGAGAAACCATATAGCCACGCC 396
 QY 62 LeuGlnAspSerAsnMetLeuPheIleThrGlyLysGlyGlyGlyLysThrCysThrGly 81
 DB 397 CTGAGGGATTCAGATCTTCTTCATTAACAGCTGGAGGGGCTGCTGATCTGCT 456
 QY 82 AlaAlaProValAlaAlaSerValAlaArgGluLeuGlyLysLeuThrValGlyValAla 101
 DB 457 GCTGCTCCAGTTGTTGCCAGATATCAAGAAAGAGCTGTTATTTACTGTGGTGTGTC 516
 QY 102 SerThrProPheArgSerGlyGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
 DB 517 ACCTATCCATTCAGTTTGCAGAGCGCTAGACGCTGTACAGCATTCGAAGCAGTACAG 576
 QY 122 GluLeuAlaLysTyrValAlaSerThrLeuIleValAlaProAsnGlnAsnLeuLeu 141
 DB 577 AAGCTGGAAGAAAGAGTGAACACACTTATTTGATTCCAATGATTAATTAATGATGTT 636
 QY 142 AlaAspLysSer 145
 DB 637 GCGGATGAAGAC 648
 RESULT 14
 LOCUS BQ012384 765 bp mRNA linear EST 22-AUG-2002
 DEFINITION OGJ1M3.yg.ab1 OG_ERCHJ lettuce serritola Lactuca sativa clone
 OGJ1M3, mRNA sequence.
 ACCESSION BQ012384
 VERSION BQ012384.1 GI:22446779
 KEYWORDS EST.
 SOURCE Lactuca sativa.
 ORGANISM Lactuca sativa.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
 Lactuca.
 1 (bases 1 to 765)
 Kozik, A., Micheltore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
 Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison,
 P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z.,
 Church, S., Jackson, L., and Bradford, K.
 Lettuce and Sunflower ESTs from the Composite Genome Project
 http://compgenome.ucdavis.edu/
 Unpublished (2002)
 CONTACT: Alexander Kozik (R.W.Micheltore)
 Department of Vegetable Crops, R.W.Micheltore Lab
 University of California at Davis (UCD)
 Asmundo Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@ucdavis.org [micheltore@vegmall.ucdavis.edu]
 belongs to contig OG_CA_Config7869, see http://cgpdb.ucdavis.edu/
 for details.
 Plate: OGJ1 row: M column: 13.
 FEATURES
 source
 1..765
 /organism="Lactuca sativa"
 /cultivar="L.serritola"
 /db_xref="taxon:4236"

```

/clone="QGJ1M13"
/clone.lib="QG_EFGHJ lettuce serricola"
/lab_host="E.coli"
/Note="Vector: pBRCDNA51AB; The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_SEQ=Not found"
BASE COUNT      208 a      153 c      208 g      196 t
ORIGIN
Alignment Scores:
Pred. No.:      1,19e-31      Length:      765
Score:          358.00      Matches:      71
Percent Similarity: 72.46%      Conservative: 29
Best Local Similarity: 51.45%      Mismatches: 38
Query Match:    40.45%      Indels:      0
DB:             14      Gaps:          0
US-09-770-509-2 (1-178) x BU012384 (1-765)
QY      41 GlysSerLysProGluLysArgSerAlaGluGlnGlnLysValAspIleGlnArg 60
DB      21 GGTGGGAATCCGATGTTGGGGGAAAGGCTCAGAGGATCAAGAACCAATTTGGAAAT 80
QY      61 MetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyThrCysThr 80
DB      81 GCCTCGAGGAGCAGATCTTGTCTCATACAGAGGAAGGAGGCTGGACAGCATCT 140
QY      81 GlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyAl 100
DB      141 GGAGCTGCTCCAGTGTGGCCAGATAGCAAAAGACAGATATCTCACCTGGTGT 200
QY      101 ValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
DB      201 GTAACCTACCCCTTCACCTTGAAGAGCGCAAAAGATCATGCCAGGCTTTGGAACTAT 260
QY      121 LysGluLeuAlaLysTyValAspThrLeuIleValValProAsnGlnAsnLeuAla 140
DB      261 GAGAACTCGAAAGAGATGTGACACACTATATCTATCCCAATGACCCCTTACTGGAT 320
QY      141 LeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyValAspAspValLeu 160
DB      321 ATTGCTGATGAGCAGACACCCCTTCAGAGATGCTTTCTTCTTGCTGATGATGATCTCG 380
QY      161 GluGlyValLysGlyValThrAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
DB      381 CAAGGCGTTCAAGGCAATTCAGATATATATACGATACCTGCGCTTGGAACGTG 434
RESULT 15
BQ869798      765 bp      mRNA      linear      EST 15-AUG-2002
LOCUS      BQ869798
DEFINITION      OGD7D11, yg. ahl QG, ABCDI lettuce salinas Lactuca sativa cDNA clone
ACCESSION      BQ869798
VERSION      BQ869798.1 GI:22255585
KEYWORDS      EST.
SOURCE      Lactuca sativa.
ORGANISM      Lactuca sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
Lactuca.
1 (bases 1 to 765)
Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison
,P., Kolman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
Church,S., Jackson,L. and Bradford,K.

```

```

TITLE      Lettuce and Sunflower ESTs from the Compositae Genome Project
JOURNAL      http://compgenomics.ucdavis.edu/
COMMENT      Unpublished (2002)
              Contact: Alexander Kozik [R.W.Michelmore]
              Department of Vegetable Crops, R.W.Michelmore Lab
              University of California at Davis (UCD)
              Asmndson Hall, UCD, Davis, CA 95616, USA
              Tel: 1-(530)-742-1742
              Fax: 1-(530)-752-9659
              Email: akozik@atgc.org [michelmoreveg@mail.ucdavis.edu]
              belongs to contig QG_CA_Config7869, see http://cgpdb.ucdavis.edu/
              for details.
              Plate: QGD7. row: D column: 11.
FEATURES
source
1..765
Location/Qualifiers
/organism="Lactuca sativa"
/cultivar="Salinas"
/db_xref="taxon:4236"
/clone="QG7D11"
/clone.lib="QG_ABCDI lettuce salinas"
/lab_host="E.coli"
/Note="Vector: pBRCDNA51AB; The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_LIB=QG_ABCDI lettuce salinas
TAG_RISUB=chemical induction
TAG_SEQ=TTGAGCCGG"
BASE COUNT      213 a      154 c      202 g      195 t      1 others
ORIGIN
Alignment Scores:
Pred. No.:      6.01e-31      Length:      765
Score:          352.00      Matches:      70
Percent Similarity: 73.33%      Conservative: 29
Best Local Similarity: 51.85%      Mismatches: 36
Query Match:    39.77%      Indels:      0
DB:             14      Gaps:          0
US-09-770-509-2 (1-178) x BQ869798 (1-765).
QY      44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
DB      8 CCGATGTTGGGGGAAACAGCTGCAGAGGATCAAGAAAGCAATCGGAATGCTTCGAG 67
QY      64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyThrCysThrGlyAlaAla 83
DB      68 GAGTACAGATCTTGTGTTATATACAGCAGAAATGGAGGTGGGACAGAGATCTGAGCTGCT 127
QY      84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValAlaSerThr 103
DB      128 CCAAGTGTGGCCCGCCAGATAGCAAAAGAACAGATATCTCACCGTTGTTGTAACCTAC 187
QY      104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
DB      188 CCGTTGACCTTTGGAAGACGCAAAAGATCAGCTCCAGGCTTTGGAAGCTATTGAGAAGCTG 247
QY      124 AlaLysTyValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
DB      248 CAAGGAATGTGGACACACTATATGTAATCCCAATATGACCGTTTACTGATTTGCTGAT 307
QY      144 LysSerThrThrMetLeuGluAlaPheArgTyValAspAspValLeuLeuGluGlyAl 163
DB      308 GACGACAGACCCCTTCAGATGCTTTCTTCTGCTGATGATGATGATCCGTCAGAGCGTT 367
QY      164 LysGlyValThrAspLeuIleValAlaArgProGlyLeuIleAsnLeu 178
DB      368 CAAGGCAATTCAGATATATATACGATACCTGCGCTTGGAACGTG 412

```

Tue Jun 3 09:14:37 2003

us-09-770-509-2.rst

Page 11

Search completed: June 2, 2003, 10:07:12
job time : 1787 secs

